

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 23, 2005, 23:04:22 ; Search time 24231 Seconds  
(without alignments)  
3379.527 Million cell updates/sec

Title: US-09-367-496C-7

Perfect score: 1690  
Sequence: 1 gccgccccctaccagagaccc.....atgaccacatgccgcgcgc 1690

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.\*

1: gb\_ba.\*

2: gb\_htg.\*

3: gb\_in.\*

4: gb\_om.\*

5: gb\_ov.\*

6: gb\_pat.\*

7: gb\_ph.\*

8: gb\_pl.\*

9: gb\_pr.\*

10: gb\_ro.\*

11: gb\_ats.\*

12: gb\_sy.\*

13: gb\_un.\*

14: gb\_vl.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1690	100.0	1690	6	AB7376 Sequence 7
2	1690	100.0	1690	6	BD070044 Use of UL
3	1689	99.9	1690	9	Y10976 H.sapiens m
4	1686.8	99.8	2699	9	AB006713 Homo sapi
5	1669.2	98.8	2727	11	BV178596 sqm10150
6	1655.8	98.0	1719	6	CQ721242 Sequence
7	1306.8	77.3	2546	10	AB006715 Mus muscu
8	1303.6	77.1	1920	6	AB7374 Sequence 5
9	1303.6	77.1	1920	6	BD070043 Use of UL
10	1303.6	77.1	1920	10	MMULIP4
11	1275	75.4	2489	10	RNUS2103
12	1067.2	63.1	2470	5	AF249294 Gallus ga
13	1066.8	63.1	1880	5	BX950824 Gallus ga
14	1038.6	61.5	2741	5	AF301551 Gallus ga
15	995.8	58.9	3884	4	BTU83278 Bos taurus
16	980.2	58.0	2947	10	RRT0AD64
17	977	57.8	4450	10	BC062955
18	970.6	57.4	1817	6	AB7370 Sequence 1
19	970.6	57.4	1817	6	BD070041 Use of UL

20	970.6	57.4	1915	10	MMULIP2	Y10339 M.musculus
21	970.6	57.4	2019	10	MMUNC33MR	X87242 M.musculus
22	965.4	57.1	1829	9	HSU17279	UI17279 Human colla
23	963.8	57.0	4459	9	AX700167	D78013 Homo sapien
24	963.8	57.0	4459	9	D78013	BC056408 Homo sapi
25	963.8	57.0	4490	9	BC056408	BC067109 Homo sapi
26	963.8	57.0	4535	9	HSN807259	BX647115 Homo sapi
27	963.8	57.0	4614	9	HSN807259	CQ861401 Sequence
28	963.8	57.0	5421	6	CQ861401	AR559299 Sequence
29	963.8	57.0	5421	6	AR559299	AX334696 Sequence
30	963.8	57.0	5421	6	AX334696	AX336393 Sequence
31	963.8	57.0	5421	6	AX336393	U97105 Homo sapien
32	963.8	57.0	5421	9	HSU97105	AF301550 Gallus ga
33	929.6	55.0	2393	5	AF301550	UI17277 Gallus gall
34	929.6	55.0	2423	5	GGU17277	AF301549 Gallus ga
35	902.4	53.4	2488	5	AF301549	U72875 Mus musculu
36	879.8	52.1	2692	10	MMU72875	AF301553 Gallus ga
37	879.6	52.0	2202	5	AF301553	BD070042 Use of UL
38	878.2	52.0	2296	6	BD070042	AB7372 Sequence 3
39	878.2	52.0	2297	6	AB7372	Y09080 M.musculus
40	878.2	52.0	2297	10	MMULIP3	BC031738 Mus muscu
41	878.2	52.0	2850	10	BC031738	AB006714 Mus muscu
42	878.2	52.0	2909	10	AB006714	BC000252 Homo sapi
43	876.8	51.9	2928	9	BC000252	BC007613 Homo sapi
44	876.8	51.9	2928	9	BC007613	CQ850591 Sequence
45	875.2	51.8	2797	6	CQ850591	

## ALIGNMENTS

### RESULT 1

LOCUS AB7376 1690 bp DNA linear PAT 22-JAN-2000  
DEFINITION Sequence 7 from Patent WO9837192.  
ACCESSION AB7376  
VERSION AB7376.1 GI:6736129  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 1690)  
AUTHORS Byk, T. and Belin, M.  
TITLE USE OF ULIP PROTEINS IN THE DIAGNOSIS AND THERAPY OF CANCER AND  
PARANEOPLASTIC NEUROLOGICAL SYNDROMES  
JOURNAL Patent: WO 9837192-A 7 27-AUG-1998;  
BYK TAMARA (CH); BELIN MARIE FRANCOISE (FR)  
FEATURES  
Location/Qualifiers  
source  
1..1690  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"

### ORIGIN

Query Match 100.0%; Score 1690; DB 6; Length 1690;  
Best Local Similarity 100.0%; Pred. No. 5.3e-265;  
Matches 1690; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 GCCGCCCTACCAGAGACCCCGAGGAGGATGTCCTTCCAGGGCAAGAAAGCATCCC 60  
Db 1 GCCGCCCTACCAGAGACCCCGAGGAGGATGTCCTTCCAGGGCAAGAAAGCATCCC 60  
Qy 61 CCGGATCAGAGTACCGCCCTTCTGATCAGAGTGGGAGGATCGTGATGACGACGATC 120  
Db 61 CCGGATCAGAGTACCGCCCTTCTGATCAGAGTGGGAGGATCGTGATGACGACGATC 120  
Qy 121 CTTTACGCTGATGTCACGTCGGAAGATGGCTTGTATAAAACAAATCGGAGAAAACCTCAT 180  
Db 121 CTTTACGCTGATGTCACGTCGGAAGATGGCTTGTATAAAACAAATCGGAGAAAACCTCAT 180  
Qy 181 CGTCCCTGGGGGCATCAAGACCATTCAGCCCGCATGATGGTCTTCTTGGTGGCGT 240  
Db 181 CGTCCCTGGGGGCATCAAGACCATTCAGCCCGCATGATGGTCTTCTTGGTGGCGT 240

Db	181	CGTCCCTGGGGGCATCAAGACCATTGACGCCACCGCGTGAATGGTCTCTCTGGTGGCGT	240
Qy	241	TGAGCTCCACAAGGCTGCAGATGCTGTCTCTGGGCATCACACCGGCTGACGACTTCTG	300
Db	241	TGAGCTCCACAAGGCTGCAGATGCTGTCTCTGGGCATGACACCGGCTGACGACTTCTG	300
Qy	301	TCAGGGCACCAAGCAGCGCTAGCAGGAGAAACCAACATGATTTGGACCAAGCTTCTCC	360
Db	301	TCAGGGCACCAAGCAGCGCTAGCAGGAGAAACCAACATGATTTGGACCAAGCTTCTCC	360
Qy	361	CGACACGGGTGTGAGCCTGTGCGGGCTACGAGCAGTGGCGGGAGCGGGCGGACAGCC	420
Db	361	CGACACGGGTGTGAGCCTGTGCGGGCTACGAGCAGTGGCGGGAGCGGGCGGACAGCC	420
Qy	421	GGCCTGTGCGACTACTCTCCCTGCACGTGGACATCACCCATGGCATGAGAGCATCAAGGA	480
Db	421	GGCCTGTGCGACTACTCTCCCTGCACGTGGACATCACCCATGGCATGAGAGCATCAAGGA	480
Qy	481	GGAGCTGGAGGCCCTGTCTCAAGGAGAAAGGTGTGAATCTCTCTCTGGTCTTCATGGCATA	540
Db	481	GGAGCTGGAGGCCCTGTCTCAAGGAGAAAGGTGTGAATCTCTCTCTGGTCTTCATGGCATA	540
Qy	541	CAAGGACCGGTGCAGTGCAGCAGCAGCAGATGTACGATCTTTCAGCATCATCTCGGGA	600
Db	541	CAAGGACCGGTGCAGTGCAGCAGCAGCAGATGTACGATCTTTCAGCATCATCTCGGGA	600
Qy	601	CCTGGGGGCTTGGCCCCAGGTGCACGCTGAGAACGGGTGTGAATCTCTCTCTGGTCTTCATGGCATA	660
Db	601	CCTGGGGGCTTGGCCCCAGGTGCACGCTGAGAACGGGTGTGAATCTCTCTCTGGTCTTCATGGCATA	660
Qy	661	CGCGTTCTCGAGCTTCGGCATCACTGCGCCCGAGGGCCACGTGCTCAGGCCACCCCGAGGA	720
Db	661	CGCGTTCTCGAGCTTCGGCATCACTGCGCCCGAGGGCCACGTGCTCAGGCCACCCCGAGGA	720
Qy	721	GGTGGAGGCTGAGCGGTGTACCGAGCTGTCAACATCGCCAAAGCAGGCAAACTGCCCGCT	780
Db	721	GGTGGAGGCTGAGCGGTGTACCGAGCTGTCAACATCGCCAAAGCAGGCAAACTGCCCGCT	780
Qy	781	GTAAGTCCACCAAGTGTATGACGAGGGGGCGGCGAGCCATCGCTCAGGCCCAAGCGCAG	840
Db	781	GTAAGTCCACCAAGTGTATGACGAGGGGGCGGCGAGCCATCGCTCAGGCCCAAGCGCAG	840
Qy	841	AGGGGTGGTCTGTCTTGGGAGGCCATCACCGCCAGCTTGGGCACCGAGCTTCACTA	900
Db	841	AGGGGTGGTCTGTCTTGGGAGGCCATCACCGCCAGCTTGGGCACCGAGCTTCACTA	900
Qy	901	CTGAGCAGAGAACTGGGCCAAGGCTGCAGCTTGTCTCATCATCCCCCTGTCAACCCAGA	960
Db	901	CTGAGCAGAGAACTGGGCCAAGGCTGCAGCTTGTCTCATCATCCCCCTGTCAACCCAGA	960
Qy	961	CCCCACACGCGACAGCACCTCACCTCTTGTCTGTCAGCGGGGACCTCCAGGTGACAGG	1020
Db	961	CCCCACACGCGACAGCACCTCACCTCTTGTCTGTCAGCGGGGACCTCCAGGTGACAGG	1020
Qy	1021	CAGCGCCCATCTGCACCTTCCACCATGCGCCAGAAAGGCTGTGGGCAAGGCAAACTTCGCGCT	1080
Db	1021	CAGCGCCCATCTGCACCTTCCACCATGCGCCAGAAAGGCTGTGGGCAAGGCAAACTTCGCGCT	1080
Qy	1081	GATCCCCGAGGCGACCAAGGCAATGAGTGTGTCAGCGGGACCTCCAGGTGACAGG	1140
Db	1081	GATCCCCGAGGCGACCAAGGCAATGAGTGTGTCAGCGGGACCTCCAGGTGACAGG	1140
Qy	1141	GGCCTCTGGGAGATGACGAGATGATGTTGTCGCGGTGACCAAGTACAAATGCTGCCAA	1200
Db	1141	GGCCTCTGGGAGATGACGAGATGATGTTGTCGCGGTGACCAAGTACAAATGCTGCCAA	1200
Qy	1201	AATCTTCAATTTTACCAAGGAGGGCGAGTGGCTGTGGGCTCTGACGCTGACCTGGT	1260
Db	1201	AATCTTCAATTTTACCAAGGAGGGCGAGTGGCTGTGGGCTCTGACGCTGACCTGGT	1260
Qy	1261	CATATGGAAACCCCAAGGCCAACAGATCATCTCTGGCCAGACCCCAAACTGGAACGTGGA	1320
Db	1261	CATATGGAAACCCCAAGGCCAACAGATCATCTCTGGCCAGACCCCAAACTGGAACGTGGA	1320

Qy	1321	GTACAACATCTTCGAGGAGTGGAGTGC CGGGAGCGCCTGCCGTGTCATTAAGTCAGGG	13380
Db	1321	GTACAACATCTTCGAGGAGTGGAGTGC CGGGAGCGCCTGCCGTGTCATTAAGTCAGGG	13380
Qy	1381	CCGAGTGGCGCTCGAGGACGGGAAGATGTTTGTACCCCGGGGGCGGGCGCTTCGTGCC	1440
Db	1381	CCGAGTGGCGCTCGAGGACGGGAAGATGTTTGTACCCCGGGGGCGGGCGCTTCGTGCC	1440
Qy	1441	TCGGAAAACATCTCCCGGACTTTGTCTACAAGAGGATCAAAGCTCGCAACAGGCTGGCGGA	1500
Db	1441	TCGGAAAACATCTCCCGGACTTTGTCTACAAGAGGATCAAAGCTCGCAACAGGCTGGCGGA	1500
Qy	1501	GATCCACGGTGTGCCCGCGGCTGTATGACGGGCGCGTCCACGAGGTGATGTGGCTGC	1560
Db	1501	GATCCACGGTGTGCCCGCGGCTGTATGACGGGCGCGTCCACGAGGTGATGTGGCTGC	1560
Qy	1561	CAGCCAGGAGGTGGGCTCCGGGCGCGGCTCTGCCAGGCAAGATCTCCGTGCTCC	1620
Db	1561	CAGCCAGGAGGTGGGCTCCGGGCGCGGCTCTGCCAGGCAAGATCTCCGTGCTCC	1620
Qy	1621	TGTGCGCAACTACATCAGTCCGGGGTTTCAAGCTTACTTGGTCTTCAGGCTGATGACCACAT	1680
Db	1621	TGTGCGCAACTACATCAGTCCGGGGTTTCAAGCTTACTTGGTCTTCAGGCTGATGACCACAT	1680
Qy	1681	CGCCCGACGC	1690
Db	1681	CGCCCGACGC	1690

RESULT 2
BD070044
LOCUS
BD070044                1690 bp       DNA            linear            PAT 27-AUG-2002
DEFINITION
Use of ULP proteins in the diagnosis and therapy of cancer and paraneoplastic neurological syndromes.

ORGANISM	<p>Homo sapiens</p> <p>Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.</p>
REFERENCE	<p>1 (bases 1 to 1690)</p>
AUTHORS	<p>Aykar, M., Belin, M.F., Honnorat, J., Kolattukudy, P., Quach, T.T., Buge, T., Sobel, A. and Aunis, D.</p>
TITLE	<p>Use of ULIP proteins in the diagnosis and therapy of cancer and paraneoplastic neurological syndromes</p>
JOURNAL	<p>Patent: JP 2001512971-A 4 28-AUG-2001;</p>
COMMENT	<p>INSTITUT NATIONAL DE LA SANTE ET DE LA RECHERCHE MEDICALE</p> <p>OS Homo sapiens (human)</p>

## ORIGIN

Query Match	100.0%;	Score 1690;	DB 6;	Length 1690;
Best Local	Similarity	100.0%;	Pred. No. 5.3e-265;	
Matches 1690;	Conservative	0;	Mismatches 0;	Indels 0; Gaps 0;
Qy	1	GCGCCCTTACCAGAGACCCCGACGACGAGATGCTCTCCAGGGCAAGAAAAGCATCCC	60	
Db	1	GCGCCCTTACCAGAGACCCCGACGACGAGATGCTCTCCAGGGCAAGAAAAGCATCCC	60	
Qy	61	CCGGATCACGAGTGACCCGCTTCTGATCAGAGTGGGAGGATCGTGAATGACCAAGTC	120	
Db	61	CCGGATCACGAGTGACCCGCTTCTGATCAGAGTGGGAGGATCGTGAATGACCAAGTC	120	
Qy	121	CTTTTACGCTGATGTCACGTGGAGATGCGTTGATATAAAACAATCGAGAAAACCTCAT	180	
Db	121	CTTTTACGCTGATGTCACGTGGAGATGCGTTGATATAAAACAATCGAGAAAACCTCAT	180	
Qy	181	CGTCCCTGGGGGCATCAAGACCAATTGACGCCACCGGCTGATGTCTCTCTGTGGCGT	240	
Db	181	CGTCCCTGGGGGCATCAAGACCAATTGACGCCACCGGCTGATGTCTCTCTGTGGCGT	240	
Qy	241	TGAGTCCACACAAGGCTGCAGATGCTCTCTGGGCATGACACCGGCTGACGACTTCTG	300	
Db	241	TGAGTCCACACAAGGCTGCAGATGCTCTCTGGGCATGACACCGGCTGACGACTTCTG	300	
Qy	301	TCAGGGACCAAGGCAGCGGTAGCAGAGAGAACCAACATGATCTTGTGAACAACGTTCCC	360	
Db	301	TCAGGGACCAAGGCAGCGGTAGCAGAGAGAACCAACATGATCTTGTGAACAACGTTCCC	360	
Qy	361	CGACACGGGTGTGAGCTGTCTGGCGGCTACGACGAGTGGCGGAGCGGGCGGACAGCGC	420	
Db	361	CGACACGGGTGTGAGCTGTCTGGCGGCTACGACGAGTGGCGGAGCGGGCGGACAGCGC	420	
Qy	421	GGCTGTCTGGGACTACTCCCTGACGCTGGACATCACCCGATGGCATCAGAGCATCAAGGA	480	
Db	421	GGCTGTCTGGGACTACTCCCTGACGCTGGACATCACCCGATGGCATCAGAGCATCAAGGA	480	
Qy	481	GGAGCTGGAGGCCCTCGTGTCAAGAGAAAGGTGTGAACTCTCTCTGTGTTTCATGGCATA	540	
Db	481	GGAGCTGGAGGCCCTCGTGTCAAGAGAAAGGTGTGAACTCTCTCTGTGTTTCATGGCATA	540	
Qy	541	CAGGACCGGTGCCAGTGCAGCGACGACGAGATGTACGAGATCTTTCAGCATCATCCGGGA	600	
Db	541	CAGGACCGGTGCCAGTGCAGCGACGACGAGATGTACGAGATCTTTCAGCATCATCCGGGA	600	
Qy	601	CCTGGGGGCTTGGCCAGGTGACGCTGAGAACGGGGAATCTGTGAGGAGGAGCAGAA	660	
Db	601	CCTGGGGGCTTGGCCAGGTGACGCTGAGAACGGGGAATCTGTGAGGAGGAGCAGAA	660	
Qy	661	CGCGTTCTGAGCTCGGCATCATCTGGCCCGGAGGGCCAGTGCTCAGCCACCCCGAGGA	720	
Db	661	CGCGTTCTGAGCTCGGCATCATCTGGCCCGGAGGGCCAGTGCTCAGCCACCCCGAGGA	720	
Qy	721	GGTGGAGGCTGAGGCGGTGTACCGAGTGTTCACCATCGCCAGGAGCAACTGCCGCT	780	
Db	721	GGTGGAGGCTGAGGCGGTGTACCGAGTGTTCACCATCGCCAGGAGCAACTGCCGCT	780	
Qy	781	GTACGTCAACAAGTGTATGAGCAAGGGGGCGGCCGACGCCATCGCTCAGGCCAAGCGCAG	840	
Db	781	GTACGTCAACAAGTGTATGAGCAAGGGGGCGGCCGACGCCATCGCTCAGGCCAAGCGCAG	840	
Qy	841	AGGGTGGTCTGTGTTTGGGAGCCCATCACCGCCAGCCTGGGGACCGGTTCACACTA	900	
Db	841	AGGGTGGTCTGTGTTTGGGAGCCCATCACCGCCAGCCTGGGGACCGGTTCACACTA	900	
Qy	901	CTGAGCAGAGAACTGGGCGCAAGGCTGAGCCTTGTGTACATCACCCTCTGTCAACCAGGA	960	
Db	901	CTGAGCAGAGAACTGGGCGCAAGGCTGAGCCTTGTGTACATCACCCTCTGTCAACCAGGA	960	
Qy	961	CCCCACCAACGGCAGACCACTCTCACTGCTGTGTGTCCAGCGGGGACCTCCAGGTGACAGG	1020	
Db	961	CCCCACCAACGGCAGACCACTCTCACTGCTGTGTGTCCAGCGGGGACCTCCAGGTGACAGG	1020	
Qy	1021	CAGCGGCCACTGCACTTTCACCATCTGCCACAGAGGCTGTGGGCAAGGACAACTTCCGCT	1080	

[illegible]

/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="ATCC No. 469253"  
/tissue\_type="retina"  
32..>1690  
/notes="mistake in reverse transcription 197..199"  
/codon\_start=1  
/product="cytosolic phosphoprotein"  
/protein\_id="CAA71872.1"  
/db\_xref="GI:1946203"  
/db\_xref="GOA:O14531"  
/db\_xref="UniProt/Swiss-Prot:O14531"  
/translation="MSFQKKISPRITSDRLIRGGRIVNDQSFYADVHVEDLIQK  
IGNLVPG1XITIDAGLMVLPQVDVHTRLOMPVLMTADDQFCQGTKEALAGTT  
MILDHVPDPDQSLAAYEOMERADSAACDYSLHVDITRMHESIKBEELALVKEG  
VNSFLVPMAYKDRQSDSYNIIFSLRDLGALAQVHAENGDIVSEEQKRLLELGT  
GPGHVLSPHEEVEAAVRAVTAQANCLYTVKMSKGAADAIAQAKRRGVVVF  
EPTTASLGTDSHYWKNWAKAAAFVTPSPNPDPTADHLTCLSSGDIQVTSAGH  
TFTTAKGAKGNFALIPETNGIIEERMSVMWEKCVASGRMDNEFVAVTSNAAKIF  
NEVPRKRVAVGSDADLVIWNPATKIISAKTNLAVEYNIPEGVRCRGAAPAVISQ  
RVALEDGRKRPVTVGAGRPVPRKTPDPVYKRIKARNRLABIHGVPRLYDGPVHEMV  
PAKPGGAPARASCPGKISVPPVRLNLSQSFSLSGQADDDHIAARR"

## ORIGIN

Query Match 99.9%; Score 1689; DB 9; Length 1690;  
Best Local Similarity 99.9%; Pred. No. 7.7e-265;  
Matches 1689; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GCCGCCCTACACGAGACCCCGAGGACAGATGTCTCCAGGGCAAGAAAGCATCCC 60  
Db 1 GCCGCCCTACACGAGACCCCGAGGACAGATGTCTCCAGGGCAAGAAAGCATCCC 60

Qy 61 CCGGATCAGAGTGACCGCCTTCTGATCAGAGTGGGAGATCGTGAATCAGCACGTC 120  
Db 61 CCGGATCAGAGTGACCGCCTTCTGATCAGAGTGGGAGATCGTGAATCAGCACGTC 120

Qy 121 CTTTATAGCTGATGTGACGTGGAAGATGGCTTGATAAAACAAATCGGAGAAACCTCAT 180  
Db 121 CTTTATAGCTGATGTGACGTGGAAGATGGCTTGATAAAACAAATCGGAGAAACCTCAT 180

Qy 181 CGTCCCTGGGGGATCAAGACCAATGACGCCACCGCTGATGGTCTCTCTGGTGGCGT 240  
Db 181 CGTCCCTGGGGGATCAGACCAATGACGCCACCGCTGATGGTCTCTCTGGTGGCGT 240

Qy 241 TGAGCTCCACAGAGCTGAGATGCTGTCTGGGATGACACCGCTGACGACTTCTG 300  
Db 241 TGAGCTCCACAGAGCTGAGATGCTGTCTGGGATGACACCGCTGACGACTTCTG 300

Qy 301 TCAGGGCACCAGCGCGCTAGCAGGAGAACACCATGATCTTGGACCACTCTTCCC 360  
Db 301 TCAGGGCACCAGCGCGCTAGCAGGAGAACACCATGATCTTGGACCACTCTTCCC 360

Qy 361 CGACAGGGTGTGAGCTGTGGCGCTACGAGCATGCGGGAGCGGCGGACAGCGC 420  
Db 361 CGACAGGGTGTGAGCTGTGGCGCTACGAGCATGCGGGAGCGGCGGACAGCGC 420

Qy 421 GGCCTGTGCGACTACTCTCCCTGACCGTGGACATCACCCGATGGCATGAGAGCATCAAGGA 480  
Db 421 GGCCTGTGCGACTACTCTCCCTGACCGTGGACATCACCCGATGGCATGAGAGCATCAAGGA 480

Qy 481 GGAAGTGGAGCGCTGTGTAAGGAGAGGGTGTGAATCTCTCTGCTCTTCAATGGCATA 540  
Db 481 GGAAGTGGAGCGCTGTGTAAGGAGAGGGTGTGAATCTCTCTGCTCTTCAATGGCATA 540

Qy 541 CAAGGACCGGTGCGAGTGACGACAGCAGCATCTTACGAGATCTTACGAGTATCATCCGGGA 600  
Db 541 CAAGGACCGGTGCGAGTGACGACAGCAGCATCTTACGAGTATCATCCGGGA 600

Qy 601 CTTGGGGGCTTGGCCAGGTGACCGCTGAGAACGGGACATCGTGGAGGAGAGCAGAA 660  
Db 601 CTTGGGGGCTTGGCCAGGTGACCGCTGAGAACGGGACATCGTGGAGGAGAGCAGAA 660

Qy 661 GCGGTTGCTGGAGCTCGGCATCACTGGCCCCGAGGGGCAAGTGTCTCAGCCACCCGAGGA 720  
Db 661 GCGGTTGCTGGAGCTCGGCATCACTGGCCCCGAGGGGCAAGTGTCTCAGCCACCCGAGGA 720

Qy 721 GGTGAGGCTGAGGCGGTGTACCGAGCTGTCAACATCGCCCAAGCAGGCAAACTGCCGCT 780  
Db 721 GGTGAGGCTGAGGCGGTGTACCGAGCTGTCAACATCGCCCAAGCAGGCAAACTGCCGCT 780

Qy 781 GTACGTCACCAAGGTGATGAGCAAGGGGGGGCGGCGCATCGCTCAGGCCAAGCGCAG 840  
Db 781 GTACGTCACCAAGGTGATGAGCAAGGGGGGGCGGCGCATCGCTCAGGCCAAGCGCAG 840

Qy 841 AGGGGTGCTGCTTTTGGGAGCCCATCACCGCAGCTGGGCAACCGAGCTTCACTA 900  
Db 841 AGGGGTGCTGCTTTTGGGAGCCCATCACCGCAGCTGGGCAACCGAGCTTCACTA 900

Qy 901 CTGAGGAAAGAACTGGGGCAAGGCTGAGCTTCTGTCATCATCACCCCTTGTCAACCCAGA 960  
Db 901 CTGAGGAAAGAACTGGGGCAAGGCTGAGCTTCTGTCATCATCACCCCTTGTCAACCCAGA 960

Qy 961 CCCACCAAGGACAGACCATCACTGCTTGTCTCCAGCGGGGACCTCCAGGTGACAG 1020  
Db 961 CCCACCAAGGACAGACCATCACTGCTTGTCTCCAGCGGGGACCTCCAGGTGACAG 1020

Qy 1021 CAGCGCCCATGCACTTCACTGCCAGAGGCTGTGGGCAAGCAAACTTTCGCGCT 1080  
Db 1021 CAGCGCCCATGCACTTCACTGCCAGAGGCTGTGGGCAAGCAAACTTTCGCGCT 1080

Qy 1081 GATCCCGGAGGCAACCAAGGCAATGAGGAGCGCATGTGATGGTCTGGGAGAAATGTGT 1140  
Db 1081 GATCCCGGAGGCAACCAAGGCAATGAGGAGCGCATGTGATGGTCTGGGAGAAATGTGT 1140

Qy 1141 GGCCTCTGGGAGATGAGAGGAATGAGTTCTGCGGGTACCAAGTACAAATGCTGCCAA 1200  
Db 1141 GGCCTCTGGGAGATGAGAGGAATGAGTTCTGCGGGTACCAAGTACAAATGCTGCCAA 1200

Qy 1201 AATCTTCAATTTTACCCAGGAAAGGGGAGTGGCTGTGGGCTCTGACGCTCACTGGT 1260  
Db 1201 AATCTTCAATTTTACCCAGGAAAGGGGAGTGGCTGTGGGCTCTGACGCTCACTGGT 1260

Qy 1261 CATATGGAACCCCAAGGCAACCAAGATCATCTCTGCAAGACCAAACTTGAAAGCTGGA 1320  
Db 1261 CATATGGAACCCCAAGGCAACCAAGATCATCTCTGCAAGACCAAACTTGAAAGCTGGA 1320

Qy 1321 GTACAACTCTTCAAGGAGTGGAGTGGCGGGAGCGCTGCGGTGATTAAGTCAGGG 1380  
Db 1321 GTACAACTCTTCAAGGAGTGGAGTGGCGGGAGCGCTGCGGTGATTAAGTCAGGG 1380

Qy 1381 CCGAGTGGCGCTGGAGGAGCGGAAAGATGTTTGTACCCCGGGGGCGGCGCTTTCGTC 1440  
Db 1381 CCGAGTGGCGCTGGAGGAGCGGAAAGATGTTTGTACCCCGGGGGCGGCGCTTTCGTC 1440

Qy 1441 TCGGAAAAATTCGCCGACTTTGTCTAAGAGGATCAAAGCTTCGCAAGCTTGGCGGA 1500  
Db 1441 TCGGAAAAATTCGCCGACTTTGTCTAAGAGGATCAAAGCTTCGCAAGCTTGGCGGA 1500

Qy 1501 GATCCAGGTGTGCGCGCTGGCTGTATGACGGGCGGCTCCACGAGTGAATGCTGCTGC 1560  
Db 1501 GATCCAGGTGTGCGCGCTGGCTGTATGACGGGCGGCTCCACGAGTGAATGCTGCTGC 1560

Qy 1561 CAAGCAGGAGTGGCGCTCCGCGCGCGCTCTGCGCCAGGCAAGATCTTCGCTGCC 1620  
Db 1561 CAAGCAGGAGTGGCGCTCCGCGCGCGCTCTGCGCCAGGCAAGATCTTCGCTGCC 1620

Qy 1621 TGTGCGCAACCTACATCAGTTCGGGGTTACGCTTATCTGGGTCTTCAGGCTGATCAACAT 1680  
Db 1621 TGTGCGCAACCTACATCAGTTCGGGGTTACGCTTATCTGGGTCTTCAGGCTGATCAACAT 1680

Qy 1681 CGCCCGACGC 1690  
Db 1681 CGCCCGACGC 1690





QY 1261 CATATGAAACCCCAAGSCCACCAGATCATCTCTGCCAAGACCCACAATCTGAACGTGGA 1320  
Db 1378 CATATGAAACCCCAAGSCCACCAGATCATCTCTGCCAAGACCCACAATCTGAACGTGGA 1437  
QY 1321 GTACAAACATCTTCAGAGGAGTGAGAGTGCAGGAGAGGCGCTTGCCTGCTGATTAAGTCAGGG 1380  
Db 1438 GTACAAACATCTTCAGAGGAGTGAGAGTGCAGGAGAGGCGCTTGCCTGCTGATTAAGTCAGGG 1497  
QY 1381 CGAGTGGCGCTGAGAGAACGGGAAGATGTTTGTACCCCGGGGCGGCGCTTGCCTGCC 1440  
Db 1498 CGAGTGGCGCTGAGAGAACGGGAAGATGTTTGTACCCCGGGGCGGCGCTTGCCTGCC 1557  
QY 1441 TCGGAAACATCTTCAGAGGAGTGAGAGTGCAGGAGAGGCGCTTGCCTGCC 1500  
Db 1558 TCGGAAACATCTTCAGAGGAGTGAGAGTGCAGGAGAGGCGCTTGCCTGCC 1617  
QY 1501 GATCCACGGTGTGCCCGTGGCTGTATGACGAGGCGCTTGCCTGCC 1560  
Db 1618 GATCCACGGTGTGCCCGTGGCTGTATGACGAGGCGCTTGCCTGCC 1677  
QY 1561 CAAGCAGGAGTGGCGCTTGCCTGCCCGGCTTGCCTGCCAGGCAAGATCTCGTGGCTCC 1620  
Db 1678 CAAGCAGGAGTGGCGCTTGCCTGCCCGGCTTGCCTGCCAGGCAAGATCTCGTGGCTCC 1737  
QY 1621 TGTGCGCAACCTACATCAGTGGGCTTGCCTGCCCGGCTTGCCTGCCAGGCAAGATCTCGTGGCTCC 1680  
Db 1738 TGTGCGCAACCTACATCAGTGGGCTTGCCTGCCCGGCTTGCCTGCCAGGCAAGATCTCGTGGCTCC 1797  
QY 1681 CGCCCGACGC 1690  
Db 1798 CGCCCGACGC 1807

RESULT 5  
LOCUS BV178596/c  
DEFINITION sqm101508 Human DNA (Sequence) Homo sapiens STS genomic, sequence tagged site.  
ACCESSION BV178596  
VERSION BV178596.1 GI:48015026  
KEYWORDS STS.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
AUTHORS Nelson,R.M., Marnellos,G., Kammerer,S., Hoyal,C.R., Shi,M.M., Cantor,C.R. and Braun,A.  
TITLE Large-Scale Validation of Single Nucleotide Polymorphisms in Gene Regions  
JOURNAL Genome Res. (2004) In press  
COMMENT Contact: Andreas Braun  
Pharmaceuticals division  
Sequenom, Inc.  
2595 John Hopkins Court, San Diego, CA 92121, USA  
Tel: 18582029018  
Fax: 18582029020  
Email: abraun@sequenom.com  
Primer A: No primer sequence submitted  
Primer B: No primer sequence submitted  
STS size: 2727.  
Location/Qualifiers  
1. .2727  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/clone\_lib="Human DNA (Sequence)"  
<1. .52727

STs  
ORIGIN

Query Match 98.8%; Score 1669.2; DB 11; Length 2727;

Best Local Similarity 99.5%; Pred. No. 1.2e-261;  
Matches 1692; Conservative 1; Mismatches 7; Indels 1; Gaps 1;  
QY 1 GCGGCCCCCTACAGAGACCCCGAGGAGGATGCTTCCAGGGCAAGAAAGCATCCC 60  
Db 2618 GCGGCCCCCTACAGAGACCCCGAGGAGGATGCTTCCAGGGCAAGAAAGCATCCC 2559  
QY 61 CCGGATCACGAGTGACCGCTTCTCATCAGAGGTGGAGGATCGTGAATGACGACGATC 120  
Db 2558 CCGGATCACGAGTGACCGCTTCTCATCAGAGGTGGAGGATCGTGAATGACGACGATC 2499  
QY 121 CTTTATACGCTGATGTGCACGTGGAGATGGCTTGAATAAACAATCGGAGAAACCTCAT 180  
Db 2498 CTTTATACGCTGATGTGCACGTGGAGATGGCTTGAATAAACAATCGGAGATTAACCTCAG 2439  
QY 181 CGTCCCTGGGGGATCAAGACCATTTGACGCCCAACCGGCTGATGGTCTTCTGGTGGCGT 240  
Db 2438 CGTCCCTGGGGGATCAAGACCATTTGACGCCCAACCGGCTGATGGTCTTCTGGTGGCGT 2379  
QY 241 TGACGTCCACACAGGCTGCAGATGCTCTCTGGGCATGACACCGGCTGACGACTTCTG 300  
Db 2378 TGACGTCCACACAGGCTGCAGATGCTCTCTGGGCATGACACCGGCTGACGACTTCTG 2319  
QY 301 TCAGGSCACCAAGGCGCTAGCAGGAGAAACCAACATGATCTTTGGACCAAGTCTTCCC 360  
Db 2318 TCAGGSCACCAAGGCGCTAGCAGGAGAAACCAACATGATCTTTGGACCAAGTCTTCCC 2259  
QY 361 CGACACGGGTGTAGGCTGCTGCGGCTTACGACAGTGGCGGAGCGGCGGACGACGC 420  
Db 2258 CGACACGGGTGTAGGCTGCTGCGGCTTACGACAGTGGCGGAGCGGCGGACGACGC 2199  
QY 421 GCGCTGCTGCGACTACTCCCTGCAGTGCATCATCCCGATGGCATGAGAGCATCAAGGA 480  
Db 2198 GCGCTGCTGCGACTACTCCCTGCAGTGCATCATCCCGATGGCATGAGAGCATCAAGGA 2139  
QY 481 GGAGCTGAGGCGCTTGGTCAAGGAGAGGCTGTGAATCTCTTCTGGTCTTTCATGGCATA 540  
Db 2138 GGAGCTGAGGCGCTTGGTCAAGGAGAGGCTGTGAATCTCTTCTGGTCTTTCATGGCATA 2079  
QY 541 CAAGGACCGGTGCGAGTGCAGGACGACGATGTACGAGATCTTCAGCATCATCCGGGA 600  
Db 2078 CAAGGACCGGTGCGAGTGCAGGACGACGATGTACGAGATCTTCAGCATCATCCGGGA 2019  
QY 601 CTTGGGGGCTTGGCCAGGTGCACGCTCAGAAACGGGGACATCGTGGAGAGGAGCAGAA 660  
Db 2018 CTTGGGGGCTTGGCCAGGTGCACGCTCAGAAACGGGGACATCGTGGAGAGGAGCAGAA 1959  
QY 661 GCGGTTGCTGGAGTCCGATCATCTGGCCCCGAGGGCCACGCTGCTCAGCCACCCGAGGA 720  
Db 1958 GCGGTTGCTGGAGTCCGATCATCTGGCCCCGAGGGCCACGCTGCTCAGCCACCCGAGGA 1899  
QY 721 GGTGGAGGCTGAGGCGGTGTACCGAGCTGTACCATCGGCNAGAGGCAAACTGCCGCT 780  
Db 1898 GGTGGAGGCTGAGGCGGTGTACCGAGCTGTACCATCGGCNAGAGGCAAACTGCCGCT 1839  
QY 781 GTACGTCAACCAAGGTGATGAGCAAGGGGCGGCGGACGCGCATCGCTCAGGCCAAGGCGAG 840  
Db 1838 GTACGTCAACCAAGGTGATGAGCAAGGGGCGGCGGACGCGCATCGCTCAGGCCAAGGCGAG 1779  
QY 841 AGGGTGGTGTGTTGGGAGCCATCATCCGCGACCTTGGGCAACCGAGTTCACACTA 900  
Db 1778 AGGGTGGTGTGTTGGGAGCCATCATCCGCGACCTTGGGCAACCGAGTTCACACTA 1719  
QY 901 CTGGAGCAAGAACTGGGGCAAGGCTGACGCTTGTGCATCATCCCCCTGTCAACCCAGA 960  
Db 1718 CTGGAGCAAGAACTGGGGCAAGGCTGACGCTTGTGCATCATCCCCCTGTGTAAACCCAGA 1659  
QY 961 CCCACACCGGAGACACCTCAGCTGCTTGTGTCTCAGCGGGGAGCTCCAGGTGACAGG 1020  
Db 1658 CCCACACCGGAGACACCTCAGCTGCTTGTGTCTCAGCGGGGAGCTCCAGGTGACAGG 1599  
QY 1021 CAGCGCCCACTGCACCTTCAACCTGCCAGAGGCTGTGGGCAAGCAAACTTTCGCGCT 1080



1021 AAGGCTGTGGGCAAGCAACATTTCCGCTGATCCCGAGGGCACCAACGGCATTTGAGGAG 1080  
 QY CGCATGTGATGTCTGGGGAATATGTGGCCCTCTGGAGATGACGAGATGATTC 1171  
 Db CGCATGTGATGTCTGGGGAATATGTGGCCCTCTGGAGATGACGAGATGATTC 1140  
 QY GTCCGGTGTACCAAGTACCAATGCTGCCAATCTTCAATTTTACCAAGGAAGGGCGA 1231  
 Db GTCCGGTGTACCAAGTACCAATGCTGCCAATCTTCAATTTTACCAAGGAAGGGCGA 1200  
 QY GTGGCTGTGGGCTCTGACGCTGACCTGCTCATATGGAACCCCAAGGCCACCAAGATCATC 1291  
 Db GTGGCTGTGGGCTCTGACGCTGACCTGCTCATATGGAACCCCAAGGCCACCAAGATCATC 1260  
 QY TCTGCCAAGACCCACCAATCTGAACGTTGGAGTACAACATCTTCAGGGAGTGGAGTCCGG 1351  
 Db TCTGCCAAGACCCACCAATCTGAACGTTGGAGTACAACATCTTCAGGGAGTGGAGTCCGG 1320  
 QY GGAGCGCTGCGCTGTGTCAATGCTCAGGGCCGAGTGGCGCTGGAGGAGTGGAGTGGTTT 1411  
 Db GGAGCGCTGCGCTGTGTCAATGCTCAGGGCCGAGTGGCGCTGGAGGAGTGGAGTGGTTT 1380  
 QY GTCAACCCCGGGCGGGCGCTTCGTCCTCGGAAACATTCGCCGACTTTTGTCTACAAG 1471  
 Db GTCAACCCCGGGCGGGCGCTTCGTCCTCGGAAACATTCGCCGACTTTTGTCTACAAG 1440  
 QY AGATCAAGCTCGCAACAGCTGGCGGAGATCAAGTGTGCCCCGTGGGCTGTATGAC 1531  
 Db AGATCAAGCTCGCAACAGCTGGCGGAGATCAAGTGTGCCCCGTGGGCTGTATGAC 1500  
 QY GGCGCGTCCACAGGTGATGTGCTCCCAAGCCAGGAGTGGCGTCCGGCCCGCG 1591  
 Db GGCGCGTCCACAGGTGATGTGCTCCCAAGCCAGGAGTGGCGTCCGGCCCGCG 1560  
 QY TCCTGCCAGGCAAGATCTCCGCTCTGTGCGCAACCTACATCAGTGGGGTTTCAGC 1651  
 Db TCCTGCCAGGCAAGATCTCCGCTCTGTGCGCAACCTACATCAGTGGGGTTTCAGC 1620  
 QY CTATCTGGGTCTAGGCTGATGACCATCGCCGAGCGC 1690  
 Db CTATCTGGGTCTAGGCTGATGACCATCGCCGAGCGC 1659  
  
 RESULT 7  
 AB006715 2546 bp mRNA linear ROD 22-AUG-1997  
 LOCUS Mus musculus mRNA for dihydropyrimidinase related protein 4,  
 DEFINITION complete cds.  
 ACCESSION AB006715  
 VERSION AB006715.1 GI:2342489  
 KEYWORDS dihydropyrimidinase related protein 4.  
 SOURCE Mus musculus (house mouse)  
 ORGANISM Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 Hamajima,N., Matsuda,K., Sakata,S., Tamaki,N., Sasaki,M. and  
 Nonaka,M.  
 1 (sites)  
 A novel gene family defined by human dihydropyrimidinase and three  
 related proteins with differential tissue distribution  
 TITLE Gene 180 (1-2), 157-163 (1996)  
 JOURNAL 97126821  
 MEDLINE 8973361  
 PUBMED  
 REFERENCE 2 (sites)  
 Hamajima,N., Kato,Y., Kouwaki,M., Wada,Y., Sasaki,M. and Nonaka,M.  
 1 (sites)  
 Novel members of dihydropyrimidinase related protein family  
 TITLE Unpublished  
 JOURNAL 3 (bases 1 to 2546)  
 REFERENCE Nonaka,M.  
 1 (sites)  
 Direct Submission  
 Submitted (19-AUG-1997) Masaru Nonaka, Nagoya City University  
 TITLE Medical School, Department of Biochemistry; 1 Kawasumi, Mizuho-cho,  
 Mizuho-ku, Nagoya, Aichi 467, Japan

(E-mail: mmonaka@med.nagoya-cu.ac.jp, Tel: 052-853-8145,  
 Fax: 052-842-3955)  
 FEATURES  
 source Location/Qualifiers  
 1. 2546  
 /organism="Mus musculus"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:10090"  
 /tissue\_type="whole embryo"  
 /dev\_stage="fetus"  
 52..1770  
 CDS  
 /codon\_start=1  
 /product="dihydropyrimidinase related protein 4"  
 /protein\_id="BAA21888.1"  
 /db\_xref="GI:2342490"  
 /translation="MSFGKKSIPIRTSDRLIKGKIVNDQSPHADLIYVEDGLIKQ  
 IGENIIVFGIKITIDAHLMLIPGVDVHTRIQMPVMTPADDFCGTGAALAGGTT  
 MILDVFDAGVSLAAVEQWRERADSAACDYSLSHVDIPRWHESTKEELRLVRDKG  
 VNSFLVFMAYKDRCCDTGQIYEIPLSLIRDLGAVQHAENGDIWVEEKRLLEQGIT  
 GPBGHVLSPBEVBAVAVTIAQKNCPLVYTKVMSGAADVMVAQAKRGVVVPG  
 EBITASLGTGDSHYMSKWKAAAFVTPSPINPDPTADHLTSLSSGDLQVTSAGH  
 TTTTAQAVGKONFTLIPGVNGIEERSVWVEKCVASGKMBENEPVATVSTNAKIF  
 NFYPRKRVAGSDADLVIMNPRATKVSASHNLNVEYNIPEGVECRGVPTVW15QG  
 RVLVDGNLLVTPGAGRFIPKTFPDVFKRIKARNRLAIEHIGVPRGLDVGPHVEML  
 PAKPGSGTQARASCSGKISVPPVRLHOSGFSLSGQADDDHIAARRTAQKIMAPPGGRS  
 NITSLN"  
 2546  
 /note="46 A nucleotides"  
 polyA\_site  
 ORIGIN  
 Query Match 77.3%; Score 1306.8; DB 10; Length 2546;  
 Best Local Similarity 86.2%; Pred. No. 9.5e-203;  
 Matches 1446; Conservative 0; Mismatches 232; Indels 0; Gaps 0;  
 QY 12 CAGAGACCCCGAGGAGGATGCTTCCAGGGCAAGAAAGCATCCCGCGATCACGA 71  
 Db |||  
 32 CAGCCCGCGGTATCAGGATGCTCTCAAGGCAAGAGAGCATCCCGGATAACGA 91  
 QY 72 GTGACCGCTCTGATCAGAGTGGGAGGATCGTGAATGACGACCATCTTTTACGCTG 131  
 Db |||  
 92 GGACCGCTCTCATCAAGGTGGGAAGATGTGAACGATGACCATCTTTTACGCTG 151  
 QY 132 ATGTGCACGTGGAAGATGGCTTGATAAAACAAATCGAGAAACCTCATCGTCCCTGGGG 191  
 Db |||  
 152 ATCTGATGTGGAAGCGCTGATTAACAAATGGAGAAATCTCATCGTCCCTGGGG 211  
 QY 192 GCATCAAGACCATTTGACGCCACGGCTGATGCTCTCTCGTGGTGGCTTACGTCACA 251  
 Db |||  
 212 GCATCAAAACCATCGATGCTCATGGCTGCTGCTGGGGAGTTGACGTTTCA 271  
 QY 252 CAAGCTGCAGATGCTCTGGGATGACACCGGCTGACGACTTCTGTCAAGGACCA 311  
 Db |||  
 272 CCGGCTGCAGATGCTCTGGATGGGATGACCCAGCTGATGATTTCTGTCAAGGACCA 331  
 QY 312 AGGACGCTAGCAGAGGAAACCAACCATGATCTTGGACCAACGCTCTTCCCGACACGGGTG 371  
 Db |||  
 332 AGCGGCTCTAGCAGCGGGGACCAACATGATTTGACCATGTTCTTCTGACGCTGGTG 391  
 QY 372 TGAGCTGTGTGGCGCTTACGAGAGTGGGGAGCGGGGGAGCAGCGGCGCTGTGCG 431  
 Db |||  
 392 TGAGCTGTGTGGCGCTTACGAGAGTGGGGAGCGGGAGCAGCAGCGGCGCTGTG 451  
 QY 432 ACTACTCCCTGCACGTGGACATCACCCGATGGCATGAGAGCATCAAGGAGGAGCTGGAG 491  
 Db |||  
 452 ACTACTCTTACATGTGGACATCTTCGCTGGCAAGAGACCAAGAGAGCTGGAG 511  
 QY 492 CCTGTGTCAGGAGAGGGTGTGAATCTCTCTCTGTCTTTCATGGCATACAAGGACCGGT 551  
 Db |||  
 512 CCTGTGTCAGGAGAGGGTGTGAATCTCTCTCTGTCTTTCATGGCATACAAGGACG 571  
 QY 552 GCAGTGCAGGACGACGAGATGTACGAGATCTTACGATCATCCCGGACCTGGGGCCT 611  
 Db |||  
 572 GCAGTGTACGAGCTGATATATGAAATCTTACGCTCATCCCGGACCTGGGGAGCTG 631

612 TGGCCAGGTGTCACGCTGAGAACGGGGACATCTGTGGAGGAGGACGACAGCGGTGCTGCG 671  
Db |||||  
632 TGGCCAGGTGTCACGCTGAGAACGGGGACATCTGTGGAGGAGGACGACAGCGGTGCTGCG 691  
Qy |||||  
672 AGCTCGGCATCACTGGGCCCGAGGGCCAGCTGCTCAGCACCCTCCAGAGAGGTGAGGCTG 731  
Db |||||  
692 AGCAAGGCATCACTGGTCTGAGGGCCATGTGCTCAGCCACCCAGAGAGGTGAGGCGG 751  
Qy |||||  
732 AGGCGGTGTCACGCTGTCACCATCGCCAGCAGGCAAACTGCCCCCTGTAGCTCACCA 791  
Db |||||  
752 AGGTGTGTACAGAGCAGTCACTTGTCCAGAGGCGCAACTGCCCACTATAGTCACCA 811  
Qy |||||  
792 AGGTGATGACAGAGGGGGCCGACGCCATCGCTCAGGCCAGCGCAGAGGGGTGTGCG 851  
Db |||||  
812 AGGTGATGACAGAGGGGTGACGTGACATGTTGTTGCCAAGCCAGCGCAGGGGTGTGCG 871  
Qy |||||  
852 TGTGTTGGGAGCCCATCAACCGCAGCTGCGGACCGAGCTTCACTACTTGGAGCAAGA 911  
Db |||||  
872 TCTTTGGGGAACCTATCACTGCCAGCTTGGGCACTGATGGCTCACTACTTGGAGCAAGA 931  
Qy |||||  
912 ACTGGCCAGAGGCTGACGCTTGTGTACATCACCCCTGTCAACCCAGAGCCCAACCAACGG 971  
Db |||||  
932 ACTGGCCAGAGGCTGACGCTTGTGTACATCACCCCTGTCAACCCAGAGCCCAACCAACGG 991  
Qy |||||  
972 CAGACCACTCACTGCTGTGTCCAGCGGGACCTCCAGGTGACAGGAGCGCCCACT 1031  
Db |||||  
992 CAGACCACTCACTGCTGTGTCCAGGTGAGGACCTCCAGGTGACAGGAGCGCCCACT 1051  
Qy |||||  
1032 GCACCTTCAACACTGCCCAGAGGCTGTGGGCAAGGACACTTCCGCTCATCCCCGAGG 1091  
Db |||||  
1052 GCACCTTCACTGCTGCCCAGAGGCTGTGGGCAAGGACACTTCCGCTCATCCCCGAGG 1111  
Qy |||||  
1092 GCACCAACGGCATTTGAGGAGCGCATGTGCTGGGAGAAATGTGGCTCTTGCGGA 1151  
Db |||||  
1112 GAGTCAACGGTATAGAGAGCGCATGTGCTGGTCTGGGAGAAATGTGGCTCTTGCGGA 1171  
Qy |||||  
1152 AGATGAGAGAGATGAGTTGTCGCGGTGACAGTACAAATGTGCCAAATCTTCAATT 1211  
Db |||||  
1172 AAATGAGAGAGATGAGTTGTCGCGGTGACAGTACAAATGTGCCAAATCTTCAATT 1231  
Qy |||||  
1212 TTTACCAAGAGAGGAGTGTGCTGCGGTGACAGTACAAATGTGCCAAATCTTCAATT 1271  
Db |||||  
1232 TTTACCAAGAGAGGAGTGTGCTGCGGTGACAGTACAAATGTGCCAAATCTTCAATT 1291  
Qy |||||  
1272 CCAAGGCCACCAAGATCATCTGTCACAGGACCCCAATCTGAACTGAGGTGACAACTCT 1331  
Db |||||  
1292 CCAGGCCACCAAGATCATCTGTCACAGGACCCCAATCTGAACTGAGGTGACAACTCT 1351  
Qy |||||  
1332 TCAGAGGAGTGAAGTGTGCGGGAGCGCTTGCCTGTGTATAGTCAAGTCAAGGCGGAGTGC 1391  
Db |||||  
1352 TTGAAGGAGTGAAGTGTGCGGGAGCGCTTGCCTGTGTATAGTCAAGTCAAGGCGGAGTGC 1411  
Qy |||||  
1392 TGAAGGAGTGAAGTGTGCGGGAGCGCTTGCCTGTGTATAGTCAAGTCAAGGCGGAGTGC 1451  
Db |||||  
1412 TGAAGGAGTGAAGTGTGCGGGAGCGCTTGCCTGTGTATAGTCAAGTCAAGGCGGAGTGC 1471  
Qy |||||  
1452 TCCGGAATTTGTCTACAGAGGATCAAGCTGCAACAGGCTGGCGGAGATCACCGGTG 1511  
Db |||||  
1472 TCCGGAATTTGTCTACAGAGGATCAAGCTGCAACAGGCTGGCGGAGATCACCGGTG 1531  
Qy |||||  
1512 TGCCCCGTGGGTGTATGACGCGGCCCTGCCAGAGGTGATGTTGCTGCCCAAGCCAGGGA 1571  
Db |||||  
1532 TGCTTCAGGCGCTGTACGACGCGGCCCTGTGCAATGAAGTGTATGTTACCTGCCAAGCCAGGAA 1591  
Qy |||||  
1572 GTGGCGCTCCGGCCCGCGCTGCTGCCAGGCAAGATCTCGGTGCTTCTGTGGCAACC 1631  
Db |||||  
1592 GTGGCAACAGGCGCGGTGCTGCTGCCAGGCAAGATCTCGGTGCTTCTGTGGCAACC 1651  
Qy |||||  
1632 TACATCAGTGGGTTTCAGGCTTATCTGGGTCTCAGGCTGATGATCAACATCGCCCGAGG 1689  
Db |||||  
1652 TGACCAAGTCTGGGTTTCAGGCTTATCTGGGTCTCAGGCTGATGATCAACATCGCCAGAGG 1709

## RESULT 8

A87374  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM

Sequence 5 from Patent WO9837192.

A87374.1 GI:6736128

Mus musculus (house mouse)

Mus musculus

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
Byk, T. and Belin, M.  
USE OF ULIP PROTEINS IN THE DIAGNOSIS AND THERAPY OF CANCER AND  
PARANEOPLASTIC NEUROLOGICAL SYNDROMES  
Patent: WO 9837192-A 5 27-AUG-1998;  
BYK TAMARA (CH); BELIN MARIE FRANCOISE (FR)

REFERENCE

AUTHORS

TITLE

JOURNAL

FEATURES

source

1. 1920

/organism="Mus musculus"

/mol\_type="unassigned DNA"

/db\_xref="taxon:10090"

ORIGIN

Query Match 77.1%; Score 1303.6; DB 6; Length 1920;

Best Local Similarity 86.1%; Pred. No. 3.3e-202;

Matches 1444; Conservative 0; Mismatches 234; Indels 0; Gaps 0;

Qy 12 CAGAGACCCCGAGGAGCGATGCTTCCAGGGCAAGAAAGCATCCCCGGATCACGA 71

Db 82 CAGCCCGCGGTGTATCAGGATGCTTCCAGGCAAGAGAGCATTTCCCGGATACGA 141

Qy 72 GTGACCGCTTCTGATCAGAGGTGGGAGGATCGTGAATGACACCACTCTTTTACGCTG 131

Db 142 GCGACCGCTTCTCATCAAGGTGGGAGATTTGAAAGATGACCACTCTTTTATGCTG 201

Qy 132 ATGTGCACTGGGAAGTGTGATATAAAACAAATCGGAGAAAACCTCATCGTCCCTGGGG 191

Db 202 ATCTGATGTGGAAGACGGTCTGATTAACAAATTCGAGAAAATCTCATCGTCCCTGGGG 261

Qy 192 GCATCAAGACCATGACGCCACCGCTGATGCTCTTCTGCTGGGTGGGTGACGTCCACA 251

Db 262 GCATCAAAACCATCGATGCTCATGGCTGATGCTGCTGGGGAGTTGACGTTCACA 321

Qy 252 CAAAGGCTGCAGATGCTGCTGCGCATGACACCGCTGACCACTTCTGTGAGGGACCA 311

Db 322 CCGGCTGCAGATGCTGCTGATGGGATGACCCAGCTGATGATTTCTGTGAGGGACCA 381

Qy 312 AGGACGCTTAGCAGGAGAACCAATGATTTGGACCACTGCTTCCCGACACCGGCTG 371

Db 382 AGGCGCTCTAGCAGGCGGACCAATGATATTGGACCATGTTTCTTCTGACGCTGGTG 441

Qy 372 TGAGCTGTGGCGGCTACGACAGTGGCGGAGCGGGGAGCAGCGCGGCTGCTGCG 431

Db 442 TGAGCTGTGGCGGCTATGAGCAGTGGCGGAGCGGAGCAGCAGCGCGGCTGCTGCTG 501

Qy 432 ACTACTCCCTGACGCTGGACATCACCCGATGGCATGAGAGATCAAGGAGAGCTGGAG 491

Db 502 ACTACTCTTACATGTTGACATCTTCTGCTGGCAAGGAGCACAAGAGAGCTGGAG 561

Qy 492 CCTGCTCAAGGAGAGGTTGAACTCTTCTGCTTCTTCTGTCATGGCATACAAGACCGGT 551

Db 562 CCTAGTCAGGAGCAAGGTTGAACTCTTCTGCTTCTTCTGTCATGGCATACAAGACCGGT 621

Qy 552 GCAGTGCAGGAGCAGCCAGATGTACAGATCTTTCAGCATCATCCGGGAGCTGGGCGCT 611

Db 622 GCAGTGTACTGACGCTCAGATATATGAAATCTTTCAGCTCATCCGGGAGCTGGGAGCTG 681

Qy 612 TGSCCCAGGTGCACGCTGAGACGGGACATCTGTGAGGAGGAGGAGGAGGAGGAGGAGG 671

Db 682 TGSCCCAGGTGCACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 741

Qy 672 AGCTCGGCATCACTGGGCCCGAGGCGCACGCTGCTCAGCCACCCCGAGGAGGTGGAGGCTG 731

```
Db 742 AGCAAGGCATCATGTCTGTAGGGCCATGTGTCTCAGCCACCCAGAGAGGTAGAGCGCG 801
Qy 732 AGCGGTGTACCGAGCTGTCCATCGCCAGCAGGCAAACTCCCGCTGTAGCTCACCA 791
Db 802 AGCGTGTACAGAGCAGTCACTATGCCAAGCAGGCCAACTGCCCACTATAGCTCACCA 861
Qy 792 AGGTGTAGCAGAGGGGGCGCGCAGCCCATCGCTCAGGCCAAGCCAGAGGGGTGGTCG 851
Db 862 AGGTGTAGCAGAGGGGTGCAGCTGTACATGTGTCCCAAGCCAGCGAGGGGGTGGTCG 921
Qy 852 TGTGTGGGAGCCCATCACCGCAGCTGTGGCCACCGAGGTTTCCACTACTCTGAGCAAGA 911
Db 922 TCTTTGGGGAACCTATCACTGCCAGCTGTGGCACTGTATGCTCACACTACTGTGAGCAGA 981
Qy 912 ACTGGGCAAGGCTGACGCTTCTGTCACTACACCCCTGTCAACCCAGACCCCAACACGG 971
Db 982 ACTGGGCAAGGCTGCAGCCTTGTGTCACTTCACTTCACTTCACTTCACTTCACTTCT 1041
Qy 972 CAGACCACTCAGCTGTCTGTCTCAGCGGGGACCTCCAGGTGACAGGCGGCCACT 1031
Db 1042 CAGACCACTCAGCTTCTGTCTGTCCAGTGGGACCTCCAGGTGACAGGCGAGTGCCCACT 1101
Qy 1032 GCACCTTCACTACTGCGCCAGAGGCTGTGGGCAAGGACAACTTTCGGGCTGATCCCGAGG 1091
Db 1102 GCACCTTCACTACTGCGCCAGAGGCTGTGGGCAAGGACAACTTTCAGTATCCCGAGG 1161
Qy 1092 GCACCAACGGCATTTAGGAGCGCATGTGATGTCTGGGAGAAATGTGTGGCTCTGTGGGA 1151
Db 1162 TAGTCAACGGTATAGAAGAGCGCATGTCTGTGTCTGTGGGAGAAATGTGTGGCTTCAAGGA 1221
Qy 1152 AGATGACGAGAAATGATTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1211
Db 1222 AAATGGAAGAGAAATGATTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1281
Qy 1212 TTTTACCAAGAGAGGCGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1271
Db 1282 TTTTACCCAGAGAGGCGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1341
Qy 1272 CCAAGGCCACCAAGATCATCTTGTCCCAAGACCCACAACTTGAAGCTGTGTGTGTGTGT 1331
Db 1342 CCAGGGCCACGAAAGTCACTTCTGCCAAGAGCCATAAAGCTGATGTGTGTGTGTGTGT 1401
Qy 1332 TCAGGGAGTGGAGTGGCGGGAGCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1391
Db 1402 TTGAAGAGTGGAGTGGCGGGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1461
Qy 1392 TGAGGAGCGGGAAGATGTTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1451
Db 1462 TGGAGGACGGAACCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1521
Qy 1452 TCCCGGACTTTGTGTCAAGAGAGTCAAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1511
Db 1522 TCCCGGACTTTGTGTCAAGAGAGTCAAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1581
Qy 1512 TGCCCGTGTGGCTGTATGACGGGCGCTGTCAAGAGTGTGTGTGTGTGTGTGTGTGTGTGT 1571
Db 1582 TGCCCGTGTGGCTGTATGACGGGCGCTGTCAAGAGTGTGTGTGTGTGTGTGTGTGTGTGT 1641
Qy 1572 GTGGCGTGTGGGCGCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1631
Db 1642 GTGGCAACAGGCGCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1701
Qy 1632 TACATCAGTGTGGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1689
Db 1702 TGCACCAAGTGTGGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1759
```

RESULT 9  
BD070043 1920 bp DNA linear PAT 27-AUG-2002  
LOCUS Use of ULIP proteins in the diagnosis and therapy of cancer and  
DEFINITION paraneoplastic neurological syndromes.

```
ACCESSION BD070043
VERSION BD070043.1 GI:22615646
KEYWORDS JP 2001512971-A/3.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE 1 (bases 1 to 1920)
AUTHORS Aguera,M., Belin,M.F., Honnorat,J., Kolattukudy,P., Quach,T.T.,
Byk,I., Sobel,A. and Aunis,D.
TITLE Use of ULIP proteins in the diagnosis and therapy of cancer and
paraneoplastic neurological syndromes
JOURNAL Patent: JP 2001512971-A 3 28-AUG-2001;
INSTITUT NATIONAL DE LA SANTE ET DE LA RECHERCHE MEDICALE
COMMENT OS Mus musculus (mouse)
PN JP 2001512971-A/3
PD 28-AUG-2001
PF 19-FEB-1998 JP 1998536335
PR 19-FEB-1997 FR 97/01961
PI MICHELE AGUERA, MARIE FRANCOISE BELIN, JEROME HONNORAT PI
, PAPPACHAN KOLATTUKUDY,
PI THAN TAM QUACH, TAMARA BYK, ANDRE SOBEL, DOMINIQUE AUNIS PC
C12N15/12, C07K14/47, C07K16/18, G01N33/574, C12Q1/68, A61K38/17, PC
A61K48/00,
PC A61K39/395
CC Use of ULIP proteins in the diagnosis and therapy of cancer
and
CC paraneoplastic neurological syndromes
FH Key Location/Qualifiers
FT source 1..1920
FT /organism='Mus musculus (mouse)'.
FEATURES
source Location/Qualifiers
1..1920
/organism='Mus musculus'
/mol_type='genomic DNA'
/db_xref='taxon:10090'
ORIGIN
Query Match 77.1%; Score 1303.6; DB 6; Length 1920;
Best Local Similarity 86.1%; Pred. No. 3.3e-202;
Matches 1444; Conservative 0; Mismatches 234; Indels 0; Gaps 0;
Qy 12 CAGAGACCCCGCCAGGACGAGTGTCTTCCAGGGCAAGAAAGCATCCCGCGATCACGA 71
Db 82 CAGCCCGCGGTGTATCAGATGTCTTCCAGGCAAGAGAGCATTCGCCGATACGA 141
Qy 72 GTGACCGCTTCTGTATCAGAGTGGAGGATCGTGAATGACGACCACTCTTTTACGCTG 131
Db 142 GCGACCGCTTCTCATCAAAAGGTGGAGATTTGGAACGATGACCACTCTTTTCATGCTG 201
Qy 132 ATGTGACGTGGAGATGCTTGTATTAACAAATCGAGAAACCTCATCTGCTCCCTGGGG 191
Db 202 ATCTGTATGTGGAAGACGGTCTGTATTAACAAATTTGGAGAAATCTCATCTCCCTGGGG 261
Qy 192 GCATCAAGACCATTTGACCGCCAGCGCTGTATGTCTCTCTGTGTGTGTGTGTGTGTGTGT 251
Db 262 GCATCAAAACCATCGATGTCTATGGCTGTATGTGTGTGTGTGTGTGTGTGTGTGTGTGT 321
Qy 252 CAAAGCTGTGAGATGCTGTCTGTGGGATGACACCGGCTGACACTTCTGTGTGTGTGTGTGT 311
Db 322 CCGGCTGTGAGATGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 381
Qy 312 AGGACGGCTAGCAGGAGGAACCACTATGTCTTGGACCACTGTCTTCCCGCAGACCGGTG 371
Db 382 AGGCGCTCTAGCAGGCGGACCACTATGTATTTGGACCACTGTGTGTGTGTGTGTGTGTGT 441
Qy 372 TGAGCTGTGTGGCGCTTACGACGTGGCGGAGCGGGCGGACAGCGCGGCTGTGTGTGTGT 431
Db 442 TGAGCTGTGTGGCGCTTATGACATGTGGCGGACGAGCAGCAGCGGCGGCTGTGTGTGT 501
Qy 432 ACTACTCCCTGTGACGTGGACATCACCCGATGGCATGAGAGCATCAAGGAGGAGCTGTGTGT 491
Db 502 ACTACTCCCTTACATGTGGACATTCCTCGCTGTGGCAGCAGAGACCAAGAGAGCTGTGTGT 561
```

Qy 492 CCCTGGTCAAGGAGAGGAGTGTGAACCTCTTCTCTGGTCTTCAATGGCATACAAAGACCGGT 551  
Db 562 CCCTAGTCAGGAGCAAAAGGTGTGAACCTCTTCTCTGGTCTTCAATGGCATACAAAGACAGGT 621  
Qy 552 GCCAGTCAGGAGCAGACGAGATGTACGAGATCTTTCAGCATCATCCGGGACCTGGGGCCT 611  
Db 622 GCCAGTGTACTGACGGGTGAGATATAGAAATCTTTCAGCCTCATCCGGGACCTGGGAGCTG 681  
Qy 612 TGGCCAGGTGTCAGCTGAGAAAGGGGACATCGTGGAGGAGGAGCAGAAAGCGGTGCTGG 671  
Db 682 TGGCCAGGTGTCAGCTGAGAAATGGGAGACATCGTGGAGGAGGAGCAGAAAGCGGTGCTGG 741  
Qy 672 AGCTCGGATCATCTGGGCCCCGAGGCGCACGCTGCTCAGCCACCCCGAGAGGTGAGGCTG 731  
Db 742 AGCAAGGCATCATCTGGTCTTGAAGGCGCATGTGCTCAGCCACCCAGAGAGGTAGAGCGG 801  
Qy 732 AGGCGGTGACCGAGCTGTACACATCGCCAGCAGCAGGCAAACTGCGCGTGTACGTACCA 791  
Db 802 AGGCTGTGTACAGAGCAGTACCAATGCGCAAGCAGGCGCAACTGCGCCACTATACGTACCA 861  
Qy 792 AGGTGATGAGCAAGGCGGCGCGCAGCGCATCGCTCAGGCCCAAGCGCAGAGGGGTGTGCG 851  
Db 862 AGGTGATGAGCAAGGCTGACGTGACATGTGTGCGCCAGCCAGCGCAGGCGGTGTGCG 921  
Qy 852 TGTGTTGGGAGCCCATCAACCGCCAGCCTGGGCAACCGAGTTCACACTTCTGGAGCAAGA 911  
Db 922 TCTTTGGGGAACCTTATCACTGCGCAGCTGGGCACTGATGGCTCACACTTCTGGAGCAAGA 981  
Qy 912 ACTGGGCGAGGCTGAGCCTTGTGTACATCAACCCCTGTCAACCCAGACCCCAACCAAGG 971  
Db 982 ACTGGGCGAAGGCTGAGCCTTGTGTACATCAACCCCTGTCAACCCAGACCCCACTACTG 1041  
Qy 972 CAGACCACTCACCTGCTGTGTCAGCGGGGACCTCCAGGTGACAGGAGCGCCCACT 1031  
Db 1042 CAGACCACTCACCTGCTGTGTCAGGTGGGACCTCCAGGTGACAGGAGTGTGCGCACT 1101  
Qy 1032 GCACCTTTCACCACTGCGCCAGAGGCTGTGGGCAAGGCAACCTTCGCGCTGATCCCCGAGG 1091  
Db 1102 GCACCTTTCACCTGCTGCGCCAGAGGCTGTGGCAAGGCAACCTTTCACCTGATCCCCGAGG 1161  
Qy 1092 GCACCAAGGATGTAGAGAGCGCATGTGATGTGTGTGGGAGAAATGTGTGGCTCTCGGA 1151  
Db 1162 TAGTCAACGGGTATAGAGAGCGCATGTGTGTGTGTGGGAGAAATGTGTGGCTCTCGGA 1221  
Qy 1152 AGATGAGACGAGATGAGTTCGTGCGGCTGACCACTACAAATGCTGCCAAATCTTCAATT 1211  
Db 1222 AAATGAGACGAGATGAGTTCGTGCGGCTGACCACTACAAATGCTGCCAAATCTTCAATT 1281  
Qy 1212 TTTACCCAGAGGAGGCGAGTGTGTGGGCTCTGAGCGCTGACCTGCTCATATGGAACC 1271  
Db 1282 TTTACCCAGAGGAGGCGGTGTGGCGCTGGGCTCTGATGCTGACCTGCTCATATGGAACC 1341  
Qy 1272 CCAAGGCCACCAAGATCATCTCTGCCAAGACCCCAAACTCTGAAGCTGGAGTACAACTCT 1331  
Db 1342 CCAGGGCCACCAAGATCATCTCTGCCAAGACCCCAAACTCTGAAGCTGGAGTACAACTCT 1401  
Qy 1332 TCAGGAGAGTGGAGTGGCGGGAGCGCTGCGGTGTGTATAGTTCAGGCGCGAGTGGCG 1391  
Db 1402 TTGAAGGAGTGGAGTGGCGGGAGTGGCGCGGTGTGTATAGTTCAGGCGCGAGTGGTGC 1461  
Qy 1392 TGAGAGGACGGGAAGATGTTGTTCACCCCGGGGGCGGCGCTTCTGCTCCCTCGGAACAT 1451  
Db 1462 TGAGAGGACGGGAACCTGCTTGTCTACTCGAGGGGCTGGCGCTTCAATCCCCGGAGAGCT 1521  
Qy 1452 TCCCGGACTTTGTCTTCAAGAGGATCAAGCTCGCAACAGCTGGCGGAGATCCAGGCTG 1511  
Db 1522 TCCCGGACTTTGTCTTCAAGAGGATCAAGCTCGCAACAGCTGGCGGAGATCCAGGCTG 1581  
Qy 1512 TGCCCCGTGGCTGTATGACGGGCGGCTTCCACGAGGTGATGCTGCTGCCAGCGGGA 1571  
Db 1582 TGCCTCAGGCGCTGTACGACGGGCGCTGTGATGAAGTGTGTTACTTCCAGCGGAG 1641

Qy 1572 GTGCGCTCCGGCCCGCGCGCTCTGCGCCAGCAAGATCTCGTGCCTCTCTGTGGCAACC 1631  
Db 1642 GTGCGCACAGCGCGCGTGCATCTGTTTCAGGCAAGATCTCAGTGCACCCCGTGGCAACC 1701  
Qy 1632 TACATCAGTCCGGGTTTTCAGGCTATCTGGGTCTCAGGCTGATGACCATCATGCCGACG 1689  
Db 1702 TGCACCACTGCGGGTTTCAGGCTATCTGGCTCTCAGGCTGACGATCACAATTGCCAGCG 1759

RESULT 10  
MMULIP4  
LOCUS  
DEFINITION M.musculus mRNA for Ulip4 protein.  
ACCESSION Y09079  
VERSION Y09079.1 GI:1915916  
KEYWORDS phosphoprotein; Ulip4 gene.  
SOURCE Mus musculus (house mouse)  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1  
AUTHORS Byk.T., Ozon,S. and Sobel,A.  
TITLE The Ulips: a family of proteins related to the axonal guidance associated unc-33 gene product  
JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 1920)  
AUTHORS Byk.T.  
TITLE Direct Submission  
JOURNAL Submitted (28-OCT-1996) T. Byk, INSERM U440, 17 rue du Per a Moulin, F- 75005 Paris, FRANCE  
COMMENT Related sequence: X87817.  
FEATURES  
Location/Qualifiers  
1..1920  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="ICR outbred strain"  
/sub\_species="Harlan Sprague Dawley"  
/db\_xref="taxon:10090"  
/tissue\_type="brain"  
/clone\_lib="lambda Zap II, Stratagene, # 936309"  
/dev\_stage="newborn"  
1..1920  
/gene="Ulip4"  
102..1820  
/gene="Ulip4"  
/codon\_start=1  
/evidence=experimental  
/product="Ulip4 protein"  
/protein\_id="CAA70299.1"  
/db\_xref="GI:1945273"  
/db\_xref="UniProt/Swiss-Prot:O35098"  
/translations="MSFQKKSIPRITSRLLIKGGKI VNDQS FHADLYVEGLIKQ IGENLIVPGGIKTIDAGLMLVPGVDVHTRLOMPVMGTPADDFCGTKAALAGGTT MILDHYFDPAGVSLAAAYBQWRGDASACDYSLSHVDIPRMHSTKEELALVRDGG VNSFLVFMAVKORCOCTDQIYEIFSLIRDLGAVAQVHAENGDI VBEQKRLLEQGIT GPGHVLSPHEEVEAEVAVRAVTIAQKNCPLVTVKMSKGAADVAQAKRGVVVFG EPTASLGTGDSGHYKSNWAKAAAFVTSPIINPDPTADHLTSLISGDLQVTSAGHC TPTAQAKGKDNFTLIPVNGCIEERMSVWEKCVASGKMDNEFVAVTSTNAKIF NTFPRKRVAGSDADLVIPWNPRATKVISAKSHNLAVENYIFEGVECRGVTPTV1SQG RVLEEDNLLVTPGAGRFIPKRTFPDFVYKIRKARNLAEIHGVPRLGYDPVHEVNL PAKPGSGTQARASCSGKISVPPVRNLHQSGFSLSGSQADHDHIAARRTAQKIMAPPGGRS NITSLG"

ORIGIN  
Query Match 77.1%; Score 1303.6; DB 10; Length 1920;  
Best Local Similarity 86.1%; Pred. No. 3.3e-202;  
Matches 1444; Conservative 0; Mismatches 234; Indels 0; Gaps 0;  
Qy 12 CAGAGACCCCGAGGAGGATGCTCTCCAGGCGCAGAAAGCATCCCGCGATCAGGA 71  
Db 82 CAGCCCGCGCGTGTATCAGGATGCTCTTCCAGGCAAGAGAGCATTTCCCGGATAACGA 141  
Qy 72 GTGACCGCCTTCTGATCAGAGGTGGGAGGATCGTGAATGACGACGAGTCTCTTTACGCTG 131



Db 142 GCGACCGCTTCTCATCAAGGTGGGAGATTGTGACGATGACAGTCTTTTCATGCTG 201  
Qy 132 ATGTGACGTTGGAGATGCTTGAATAAACAATCGAGAGAAACCTCATGTCCTCGGG 191  
Db 202 ATCTGTATGTGGAAGACGGTCTGATTAAACAAATTGGAGAAATCTCATGTCCTCGGG 261  
Qy 192 GCATCAAGACCATTTGACGCCCAAGGCTGATGCTCTTCTGTTGGGTTGAGTCCACA 251  
Db 262 GCATCAAAACCATTCGATGCTCATGCGCTGATGCTGCTGCGGAGTTGACGTTCA 321  
Qy 252 CAAAGGCTGACAGTGCCTCTCTCGGCATGACACCGCTGACGACTTCTGTTCAGGACCA 311  
Db 322 CCGGCTGACAGTGCCTGTGATGCGCATGACCCAGCTGATGATTTCTGTTCAGGACCA 381  
Qy 312 AGCAGCGCTAGCAGAGGAAACCAACATGATCTTTGACACGCTCTTCCCGCAGCACGGGTG 371  
Db 382 AGCGGCTTAGCAGCGCGGACCAACATGATATTTGGACCATGTTGTTTCTGACGCTGGTG 441  
Qy 372 TGAGCCTGCTGGCGCTTACGACAGTGGCGGAGCGGGGACAGCGCGGCTGCTGGG 431  
Db 442 TGAGCCTGCTGGCAGCTATGAGCAGTGGCGGACGAGCAGCGCGGCTGCTGGT 501  
Qy 432 ACTACTCCTGACGCTGACATCAACCGATGCGATGAGCATCAAGGAGGAGCTGGAGG 491  
Db 502 ACTACTCCTTACATGTGGACATCTCTCGCTGGACGAGACCAACAGAGAGCTGGAGG 561  
Qy 492 CCTGTGTTCAAGAGAGAGGTGTGAATCTCTCTCTGTTTCATGCGCATACAAGACCGGT 551  
Db 562 CCTAGTTCAGGACCAAGGTGTGAATCTCTCTCTGTTTCATGCGCATACAAGACAGGT 621  
Qy 552 GCAGTGCAGGACAGCAGATGTACGAGATCTTTCAGCATCATCCGGGACCTGGGGCCT 611  
Db 622 GCCAGTGTACTGACGCTGACATATGAAATCTTTCAGCCTCATCCGGGACCTGGGAGCTG 681  
Qy 612 TGCGCCAGTGTGACGCTGAGAACCGGGACATCTGTGAGGAGGAGCAGAACGGTGTGCTGG 671  
Db 682 TGCGCCAGTGTGACGCGAGAAATGCGGACATCTGTGAGGAGGAGCAGAACGGTGTGCTGG 741  
Qy 672 AGCTGCGCATCATTTGGCCCCGAGGCGCACTGTCTGACGCCACCGCGAGAGGTGAGGCTG 731  
Db 742 AGCAAGGACATCACTGGTCTCTGAGGCGCATGTGCTCAGCCACCCAGAGAGGTAGAGCGCG 801  
Qy 732 AGCGGTTGACGAGCTGTACCATCGCCAGCAGGCAAACTGCCGCTGATGTCACCA 791  
Db 802 AGGCTGTGTACAGAGCAGTCAACATTTGCCAAGCAGGCGCAATGCCCACTATACGTACCA 861  
Qy 792 AGGTGATGAGCAAGGGGCGCGCAGCCCATCGCTCAGGCCAAGCGCAGAGGGTGTGCTG 851  
Db 862 AGGTGATGAGCAAGGGTGCAGCTGACATGTTGCGCCAAAGCCAGCGGAGGGGTGTGCTG 921  
Qy 852 TGTGTTGGGAGCCCATCACCGCAGCTGGGCAACCGAGCGTTTCACTACTGAGGACAGA 911  
Db 922 TCTTTGGGGAACCTATCACTGCCAGCTGGGCACTGATGCTCACACTACTGAGGACAGA 981  
Qy 912 ACTGGGCAAGGCTGAGCCTTGTGTCATCATCAACCCCTGTCAACCCAGACCCACACCGG 971  
Db 982 ACTGGGCAAGGCTGAGCCTTGTGTCATCTTCAACCCCTATCAACCCGAGCCCTACTG 1041  
Qy 972 CAGACCACTTCACTGCTGTCTGTCAGCGGGGACCTCCAGGTGACAGGAGGCGCCCACT 1031  
Db 1042 CAGACCACTTCACTGCTGTCTGTCAGTGGGACCTCCAGGTGACAGGAGTGGCCACT 1101  
Qy 1032 GCACCTTTCACACTGCGCCAGAGCTGTGGGCAAGGACAACTTCGCGCTGATCCCGAGG 1091  
Db 1102 GCACCTTTCACACTGCGCCAGAGGCTGTGGCAAGGACAACTTCACACTGATCCCGAGG 1161  
Qy 1092 GCACCAAGGCTTGAAGAGGCGCATGTCGATGTTCTGGGAGAAATGTGTCCTCTGGGA 1151  
Db 1162 TAGTCAACGGTATAGAAGAGCGCATGTCTGTGCTCTGGGAGAAATGTGTGGCTTCAGGA 1221  
Qy 1152 AGATGAGCAGAGATGAGTGTGTCGCGGTGACCAAGTACAAATGCTGCCAAAATCTTCAATT 1211

Db 1222 AAATGACGAGAATGAGTTGTTGCTTGACGACCAAAATGCTGCCAAAATCTTCAATT 1281  
Qy 1212 TTATCCCAAGGAAGGGCGAGTGGCTGTGGGCTCTGACGCTGACCTGGTTCATATGGAAC 1271  
Db 1282 TTATCCCAAGGAAGGGCGAGTGGCTGTGGGCTCTGATGCTGACCTGGTTCATATGGAAC 1341  
Qy 1272 CCAAGGCCACCAAGATCATCTCTGCGCAAGACCAAACTCTGAAACGTGGAGTACAAATCT 1331  
Db 1342 CCAAGGCCACCAAGATCATCTCTGCGCAAGAGCAAACTCTGAAATGTAGAGTACAAATCT 1401  
Qy 1332 TCAGAGGAGTGGAGTGCCTGGGAGCGCTGCGTGTGTCTAATAGTCAAGGCGGAGTGGCGC 1391  
Db 1402 TTGAAGGAGTGGAGTGCCTGGGAGTGCCTGCAAGGAGTGCCTGCAAGGAGTGGTGC 1461  
Qy 1392 TGAAGGACGGGAGATGTTTGTCAACCGGGGGCGGCGCTTCCCTCCCGGAAAACAT 1451  
Db 1462 TGAAGGACGGGAAACCTGCTTGTCACTCCAGGGGCTGGCGCTTCATTTCCCGGAAACGT 1521  
Qy 1452 TCCCGGACTTGTCTACAAAGAGATCAAAAGCTCGCAACAGGCTGGCGGAGATCCAGGTG 1511  
Db 1522 TCCCGGACTTGTCTATAGAGGATAAAGCTTCGCAACAGGCTAGCAGAGATCCAGGTG 1581  
Qy 1512 TGCCCCGTGGGCTGTATGACGCGGCCCTCCACGAGGTGATGTTGCTGCAAGGCGAGGGA 1571  
Db 1582 TGCCCTCGAGGCTGTACGACGCGGCTGTGCATGAAGTGTATCTGCTGCAAGGCGAGGAA 1641  
Qy 1572 GTGGCGCTCGGCGCGGCGGCTGCTGCCAGGCGCAAGATCTCCGTGCTCTGTGCGCAAC 1631  
Db 1642 GTGGCACACAGGCGCGTGTCTGTTTCAGGCAAGATCTCAGTGGCCACCGTGGCGAAC 1701  
Qy 1632 TACATCAGTTCGGGTTTTCAGCTATCTGCGTCTCAGGCTGATGACCAATCGCCCGACG 1689  
Db 1702 TGACCAAGTTCGGGTTTTCAGCTATCTGCGTCTCAGGCTGATGATCATGTCAGTGGCAGAG 1759

RESULT 11  
RNU52103 2489 bp mRNA linear ROD 27-JUL-2001  
LOCUS Rattus norvegicus rCRMP-3 mRNA, partial cds.  
DEFINITION U52103  
ACCESSION U52103.1 GI:1399539  
VERSION  
KEYWORDS  
SOURCE Rattus norvegicus (Norway rat)  
ORGANISM  
Rattus norvegicus  
Bukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
Rattus.  
REFERENCE 1 (bases 1 to 2489)  
AUTHORS Wang L.H. and Strittmatter S.M.  
TITLE A family of rat CRMP genes is differentially expressed in the nervous system  
J. Neurosci. 16 (19), 6197-6207 (1996)  
MEDLINE 96424532  
PUBMED 8815901  
REFERENCE 2 (bases 1 to 2489)  
AUTHORS Strittmatter S.M.  
TITLE Direct Submission  
JOURNAL Submitted (22-MAR-1996) Stephen M. Strittmatter, Neurology, Yale University School of Medicine, 333 Cedar Street, New Haven, CT 06520, USA

FEATURES  
source 1..2489  
/organism="Rattus norvegicus"  
/mol\_type="mRNA"  
/db\_xref="taxon:10116"  
/tissue\_type="brain"  
/dev\_stage="adult"  
<1..1696  
/note="similar to 62 kDa chick collapsin response mediator protein"  
/codon\_start=2  
/product="rCRMP-3"  
/protein\_id="AAB03281.1"

CDS

		/db_xref="GI:1399540"			
		translation="IPRITSDRLIKGGKIIVNDQSFADLYVEDGLIKQIGENLIYP GGTIDAHGLMVLPGVQVHTRLQMPVLGTPADPCQGTQAGTGMILHDHVP DAGVLLAAVEQWRERADSAACDYSLSHVDI PRWHBETKEELRALVRDKVNSFLVPM AYDRCCOTDGOIYBLSIRLDGAQAHAENGDI VEERQKRLLEQIITGPBHVLS HPREVEAYTRAVTIKQANCPYIYTKMSKGNADVAQKRGVVFPBEPITASLG TDSHYWKNWAKAAAFVTPSPINPDFTADHLTSLUSSGDLQVTSAAHCTFTTAQKA VGKDNFTLIPEGINGIERMSVWVEKCVASGRMDNEFVAVTSTNAKI FNFPYRGR VAVGSDADLIWPRATKVI SAKSHNLNVEYNIPEGVCEGRMPTVVISQGRVLEGN LLVTPGAGRIPKRTFPDFVKRI KARNRLAEIHGVPRGLYDGPVHEVMLPAKPGSGT QARASCPGKISVPPVRNLHOSGFSLSQSDHIAARTAKIMAPPGRSNIITLS"			
		75.4%; Score 1275; DB 10; Length 2489; Best Local Similarity 86.2%; Pred. No. 1.4e-197; Matches 1410; Conservative 0; Mismatches 225; Indels 0; Gaps 0;			
		ORIGIN			
Qy	56	ATCCCCGGATCAGAGTGACCGCTTCTGATCAGAGGTGGGAGGATCGTGAATGACGAC	115	896	CACCTACTGGAGCAAGAACTGGGCCAAGGCTGACGCTTCTCACAATCACCCCTGTCAAC
Db	2	ATCCCCGGATAACGAGTGACCGCTTCTCATCAAGGTTGGGAAGATTGTGAATGATGAC	61	842	CACCTACTGGAGCAAGAACTGGGCCAAGGCTGACGCTTCTCACAATCACCCCTGTCAAC
Qy	116	CAGTCCCTTTACGCTGATGTCACGTGCAAGATGGCTTGATATAAAACAAATCGGAGAAAC	175	956	CCAGACCCACCAAGGCGAGACCACTCAGCTGCTTCTGCTCCAGCGGGGACCTCCAGGTG
Db	62	CAGTCCCTTTATGCGGATCTCTAGTGGAGGACGGCTGATTAACAAATTTGGGAAAT	121	902	CCGAGCCCTACTACTGTCAGACCACTCACCCTCTCTGCTGTCAGTGGGGACCTCCAGGTG
Qy	176	CTCATGCTCCCTGGGGGATCAAGACCATTAAGACCCCAAGCTGATGGTCTTCTCTGGT	235	1016	ACAGGCGAGCCGCACTGTCACCTTTCCACCTGCCCCAGAAAGGCTGTGGGCAAGGACAACTTC
Db	122	CTTATGCTCCCTGGGGGAAATCAAAACCATCGATGCTCATGGTCTGATGGTACTGCTGGG	181	962	ACAGGCGAGTGCACACTGTCACCTTCACTACTGCCCAAGAGGCTGTGGCAAGATAACTTC
Qy	236	GGCGTTGACGTCCACACAGGCTGCAGATGCTGCTGCTGGGCATGACACCGCTGACGAC	295	1076	GCCTGATTCCTCCAGGAGCAACACGGCATTTGAGAGGCGCATGTCGATGGTCTGGGAGAAA
Db	182	GGAGTTGACGTCCACACCGGCTGCAGATGCTGCTGCTGGGCATGACACCGCTGACGAC	241	1022	ACGCTGATTCCTCCAGGAGCAACACGGCATTTGAGAGGCGCATGTCGCTGGTCTGGGAGAAA
Qy	296	TTCTGTCAAGGACCAAGGACGCTTAGCAGGAGGAAACACCATGATCTTGGACCAAGTC	355	1136	TGCTGCTGCTTCAAAATTTTACCAAGGAGGCGGAGTGGCTGTGGCTCTGACGCTGAC
Db	242	TTCTGTCAAGGACCAAGGACGCTTAGCAGGAGGAAACACCATGATCTTGGACCAAGTC	301	1082	TGCTGCTGCTTCAAAATTTTACCAAGGAGTGGAGTGGCTGTGGCTGACCAAAATGCT
Qy	356	TTCCCGACACCGGCTGTGAGCTGCTGCGGCTTACGAGCAGTGCGGAGCGGCGGAC	415	1196	GCCTGATTCCTCCAGGAGCAACACGGCATTTGAGAGGCGCATGTCGATGGTCTGGGAGAAA
Db	302	TTTCTGATGCTGGCGTGAGCTGCTGCGGCTTACGAGCAGTGCGGAGCGGCGGACGAC	361	1142	GCCTGATTCCTCCAGGAGCAACACGGCATTTGAGAGGCGCATGTCGATGGTCTGGGAGAAA
Qy	416	AGCGGCGCTGCTGGGATCTCCTCGATGCACTGAGACATACCGATGGCATGAGAGATC	475	1256	CTGGTCAATGGAACCCCAAGGCCACCAAGATCATCTCTGCCAAGACCAATCTGAAC
Db	362	AGTGACGCTGCTGAGCTACTCTTACACGTGACATTTCCCGCTGGCATGAGAGACCC	421	1202	CTGGTCAATGGAACCCCAAGGCCACCAAGATCATCTCTGCCAAGACCAATCTGAAC
Qy	476	AAGGAGGAGCTGGAGGCGCTGCTCAAGAGAGGCTGTAACCTCTTCTGCTTCTCATG	535	1316	GTGGAGTACCAATCTTCCGAGGAGTGGAGTGGCGGGAGCGCTGCGGTGCTCATAGT
Db	422	AAAGAGGAGCTGGAGGCGCTGCTCAAGAGAGGCTGTAACCTCTTCTGCTTCTCATG	481	1262	GTAGAGTACCAATCTTCCGAGGAGTGGAGTGGCGGGAGCGCTGCGGTGCTCATAGT
Qy	536	GCATACAGGACCGGTGCGGATGTCAGGACGACGACGATGTACGAGATCTTTCAGCATCATC	595	1376	CAGGCGCGAGTGGCGCTGGAGGAGCGGAGATGTTGTACCCCGGGGGCGGCGCTTC
Db	482	GCTTACAGGACCGGTGCGGATGTCAGGACGACGACGATGTACGAGATCTTTCAGCATCATC	541	1322	CAGGCGCGAGTGGCGCTGGAGGAGCGGAGATGTTGTACCCCGGGGGCGGCGCTTC
Qy	596	CGGACCTGGGGGCTTGGCGGAGGCTGACCTGAGAAACGGGAGCATGCTGGAGGAGGAG	655	1436	GTCCCTCGGAAAAACATTCCTCGGACTTTGTCTACAAGAGGATCAAGCTTCGCAACAGGCTG
Db	542	CGGACCTGGGGGCTTGGCGGAGGCTGACCTGAGAAACGGGAGCATGCTGGAGGAGGAG	601	1382	ATTCCCGGGAAGACATTTCCAGACTTTGTCTATAAGAGGATAAAGGCTCGTAAACAGGCTA
Qy	656	CAGAGCGGTGCTGGAGCTCGGATCACTGCGCCCGAGGGCGACGCTGCTCAGCCACCC	715	1496	CGGAGATTCACGCTGCTGCGGCTGATGACGGGCGGCTCCAGGAGTGTGCTG
Db	602	CAGAGCGGTGCTGGAGCTCGGATCACTGCGCCCGAGGGCGACGCTGCTCAGCCACCC	661	1442	GCAGAGATTCACGCTGCTGCGGCTGATGATGATGGGCTGTCCACGAGTGTGTTA
Qy	716	GAGGAGTGGAGGCTGAGGCGGCTGACGAGCTGTACCATCGCAAGAGGAGGCAAACTGC	775	1556	CTGCGCAAGCAGGAGTGGCGCTCGCGCCCGCGCTGCTGCGCAGGCAAGATCTCCGTG
Db	662	GAGGAGTGGAGGCTGAGGCGGCTGACGAGCTGTACCATCGCAAGAGGAGGCAAACTGC	721	1502	CTGCGCAAGCAGGAGTGGCGCTCGCGCCCGCGCTGCTGCGCAGGCAAGATCTCAGTG
Qy	776	CCGCTGTACGTACCAAGGTGATGAGCAAGGGGCGGCGGACGATGCTGTCAGGGCCAAG	835	1616	CTTCTGTGCGCAACCTTACATCAGTCCGGGTTTCAAGCTATCTGGTCTCAGGCTGATGAC
Db	722	CCACTGTATATACCAAGGTGATGAGCAAGGGTGGGCTGATATGTTGGCCCAAGCCAA	781	1562	CCACCGTGGGCAACCTGACACAGTCAAGGTTTCAAGGCTATCTGGCTCTCAGGCTGACGAT
Qy	836	CGCAGAGGGTGGTGGTGGGAGCCCATCACCGCCAGGCTGGGCGGACCGGTTCA	895	1676	CACATCGCCGACGCG 1690
Db	782	CGCAGAGGGTGGTGGTGGGAGCCCATCACCGCCAGGCTGGGCGGACCGGTTCA	841	1622	CACATTCAGGAGCGC 1636
		RESULT 12			
		AP249294			
		LOCUS			
		Gallus gallus collapsin response mediator protein-3B (CRMP3B) mRNA,			
		complete cds.			
		AP249294			
		VERSION			
		AP249294.1 GI:11559529			
		KEYWORDS			
		SOURCE			
		ORGANISM			
		Gallus gallus (chicken)			
		Gallus gallus			
		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
		Archosauria; Aves; Neognathae; Galliformes; Phasianidae;			
		Phasianinae; Gallus.			
		1 (bases 1 to 2470).			
		REFERENCE			
		AUTHORS			
		Fukada, M., Watakabe, I., Yuasa-Kawada, J., Kawachi, H., Kuroiwa, A.,			
		Matsuda, Y. and Noda, M.			

TITLE	Molecular characterization of CRMP5, a novel member of the collapse response mediator protein family	Db	309	GAGGACACGATGATAATGGACCATGCTGCTGCCGAGGCTGGAGAGGCTGCTGGCTG	368
JOURNAL	J. Biol. Chem. 275 (48), 37957-37965 (2000)	Qy	387	CTTACGAGCAGTGGCGGGAGCGGGGAGCAGCGCGGCTGCTGCGACTACTCTCCCTGACG	446
MEDLINE	20545548	Db	369	CTTACGAGCAGTGGCGGGAGCGGGGAGCAGCGCGGCTGCTGCGACTACTCTCCCTGACG	428
PUBMED	10956643	Qy	447	TGGACATCACCCGATGGCATGAGAGCATCAAGGAGGAGCTGGAGGCGCTGGTCAAGGAG	506
REFERENCE	2 (bases 1 to 2470)	Db	429	TGACATCCCCCGCTGGCATGAGAGCTGCGGAGGAGCTGGAGGCGCTGGTGAAGGACA	488
AUTHORS	Yuasa-Kawada, J., Suzuki, R., Kano, F., Ohkawara, T., Murata, M. and Noda, M.	Qy	507	AGGCTGTGAATCTCTTCTCTGCTTTCATGGCATACAAAGACCGGTCGCCAGTGCAGGACA	566
TITLE	Axonal morphogenesis controlled by antagonistic roles of two CRMP subtypes in microtubule organization	Db	489	AGGCTGTGAATCTCTTCTCTGCTTTCATGGCATACAAAGACCGAGCTGCAGTGCACGAG	548
JOURNAL	Eur. J. Neurosci. 17 (11), 2329-2343 (2003)	Qy	567	GCAGATGTCAGAGATCTTTCAGCATCATCCGGGAGCTGGGGGCTGGCCAGGTGCAGC	626
MEDLINE	22699123	Db	549	CCAGATGTATGAAATATTTCTGATATTTTCGAGACCTGGGGGCTATAGCCCAAGTGCACG	608
PUBMED	12814366	Qy	627	CTGAGAACGGGACATCGTGGAGGAGGAGCAGGAGGCTTGTCTGGAGCTCGGCATCACTG	686
REFERENCE	3 (bases 1 to 2470)	Db	609	CTGAAACCGGGGACATCATCGAGGAGGAGCAGGAGGAGGCTGCTGGACATTTGGGATCACTG	668
AUTHORS	Yuasa-Kawada, J., Watakabe, I. and Noda, M.	Qy	687	GCCCCAGGGCCACATGCTCAGCCACCCCGAGGAGGTGGAGGCTGAGGCGGTGTACCGAG	746
TITLE	Direct Submission	Db	669	GGCCAGAGGGGACATGCTCCTGAGCGGCTTGAAGGAGGTTCAGCAGAGGCGGTGTACCGTG	728
JOURNAL	Submitted (27-MAR-2000) Division of Molecular Neurobiology,	Qy	747	CTGTACCATTCGCAAGCAGGAGGAACTGCCCGCTGTACCTACCAAGGTGATGAGCAAGG	806
MEDLINE	National Institute for Basic Biology, Myodaiji-cho, Okazaki, Aichi	Db	729	CAATTACAATAGCAAAACAAGCAACTGCCCTTGTACGTCACCAAAATAATGAGCAAA	788
PUBMED	444-8585, Japan	Qy	807	GGGGGCGGACGCTCATCGCTCAGGCCAAGCGAGGAGGCTGCTGGTGGGAGGCCA	866
REFERENCE	1. .2470	Db	789	GTGACGACATGTGTAGTCTCAAGCAAGAGAGGAGGCTGTGGTGTACGGGAGGCCA	848
AUTHORS	Yuasa-Kawada, J., Watakabe, I. and Noda, M.	Qy	867	TCACCCCGGACCTGGGACACGACGCTTACACTACTGAGCAAGAACTGGGCCAAGGCTG	926
TITLE	translation="MENPKRSRGPRGTSEQLIKGGKIVNDQSFYADIYVEDLIKQ	Db	849	TCACCTGCCAGCTGGGTGTGACGGCTCGCACTACTTGGAGCAAGAACTGGGCCAAGGCTG	908
JOURNAL	IGNLAVPSQVTVAYGOLVPGIDVHTLQMAVMGASADDFYQGTAAALAGTT	Qy	927	CAGCTTGTGTCACATCACCCCTGTCAACCCAGACCCACACGCGGAGACCACTCACTACCT	986
MEDLINE	MINDVLPRAGTSLLAAYEOMRORADSRACDYALHIDIPRWHSERLEALVADKDG	Db	909	CAGCTTGTGTCACCTCACCCCTGTCAACCCCGGCTCAGCCCGGACCCACACGCGGAGCGCTCTCTCT	968
PUBMED	VNSFLVPMAYKORLOCTDAQYEIFCLIRDLAIAQVHAENGDIIEBEQKRLDIDGIT	Qy	987	GCTTGTGTCGAGGGGAGCTCCAGGTGACGAGGAGCGCCACTGCACCTTCCACCACTG	1046
REFERENCE	GPGHVLSPREEBAVYNAITIAQANCLPVTKIMSKSAADVVAQAQRKGTVVYG	Db	969	CCCTGCTGCTGTGGGGACCTGACAGTGGGAGGAGTGTCTCACTGACGCTTCACTACCG	1028
AUTHORS	EPITASLGADGSHYFNKNWAKAAAFVTPSPDPPTPERLSLSDGLQVAGSAHC	Qy	1047	CCGAGAGGCTGTGGGCAAGGACAACTTCGCGCTGATCCCGAGGCGACCAAGCGCATTTG	1106
TITLE	TFTTAQKAGNKTILPEGTNGIERMAIIMBKCVVSGWQNDENPVAVTSTNAKIF	Db	1029	CCGAGAGGCTGTGGGCAAGGACAACTTCACCTCTCATCCCGAGGAGGACGAATGGCATCG	1088
JOURNAL	NCVPRKRIAVSGDADLVNPRATKVISAKTNLAVNYNIFEGTECHGAPAVVISQ	Qy	1107	AGGAGCGCATGTGATGCTGGGAGAAATGTGTGCGCTCTGGGAGATGGAAGATG	1166
MEDLINE	KVLELGNLPTVSGSGFRPKTFPDVYKRIKARNLRAVHGVPRGLVDGPVHDVL	Db	1089	AGGAGCGCATGTGCAATACTGGGAGAAATGTGTGCGCTCTGGGAGATGGAAGATG	1148
PUBMED	STKAVPTIAASRIAACGKVPVAVRNHQSGLSGLSQDADHVARIAQKIMAPGG	Qy	1167	AGTTCTGCGGTGACAGTACAAATGCTGCGCAAAATCTTCAATTTTACCCAGGAGG	1226
REFERENCE	RSNITSLS"	Db	1149	ATTTCTGCGGTGACAGCAGCTTAATGCTGCGCAAAATCTTAACTGCTACCCCGAGG	1208
AUTHORS		Qy	1227	GGCGAGTGGCTGTGGGCTGTGACGCTGACCTGTGTATATGAAACCCCAAGGCCCAAGA	1286
TITLE		Db	1209	GGCGAATCGCGGTGGGCTGTGATGACAGCTTGTGTGTGGAACCCAGGGCTACAAAG	1268
JOURNAL		Qy	1287	TCATCTCTGCAAGA CCAATCTGAA CGTGGAGTACAACTCTTCGAGGGAGTGGAGT	1346
MEDLINE		Db	1269	TCATCTCGGCAAAA CCAATTTGAA TGTGGAGTACAACTATTTTGAAGGACGCGAGT	1328
PUBMED		Qy	1347	GCAGGGAGCGCTCGCGGTGTCATTAAGTCAGGCGGAGTGGCGCTGGAGGAGGAGGA	1406
REFERENCE		Db	1329	GTACCGCGCTCTCTGCTGTGTGTATCAGTCAGGCGCAAGTCTTGTGAGGACGGGAAC	1388
AUTHORS		Qy	1407	TGTTTGTCAACCCCGGGGCGGCGCTGTCTCCCTCGGAAAAATTTCCCGGACTTTGTCT	1466

## ORIGIN

Query Match	63.1%;	Score 1067.2;	DB 5;	Length 2470;
Best Local Similarity	77.9%;	Pred. No. 8.3e-164;		
Matches 1301;	Conservative 0;	Mismatches 363;	Indels 6;	Gaps 1;
Qy	27	GCAGGATGCTCTCCAGGGCAAGAAAGCATCCCCGGATCAGAGTGACGGCTTCTGA	86	
Db	9	GGAGATGTCGAACCCGGGCAAGAGGAGCGGCCCGCGGACGAGCAGCGTGTGA	68	
Qy	87	TCAGAGTGGAGGATCGTGAATGACGACGACGCTCTTTTACGTGTGTCACGCTGAAG	146	
Db	69	TCAAGGGGGGAAAATCGTCAACGACGACGACGCTGTTTATGCCGACATTTATGTGGAAG	128	
Qy	147	ATGGCTTGATAAAACAAATCGGAGAAACCTCATCGTCCCTGGGGGCATCAAGACCAATTG	206	
Db	129	ATGGTTGATAAAACAGATTTGGAGAGAACTGGCTGTTCACGCGGGTGAGGACGGTGG	188	
Qy	207	ACGCCCAACGCGCTGATGGTCTTCTGCTGGTGGCGTTGACGTCCACACAAGGCTGCAGATGC	266	
Db	189	ATGCTTACGGGCGACGTGGTGTGTCGGCGCGGCATCGATGTGTGCAACGCGGCTGCAGATGG	248	
Qy	267	CTGTCTGGGCATGACACCGGCTGACCACTTCTGTACAGGGCACCAAGGACGCGCTAGCAG	326	
Db	249	CTGTGATGGGGATGGCTCTGCTGATGACTTCTACCAAGGACGAGGGCGGCCCTGGCGG	308	
Qy	327	GAGGAACCAACCATGATCTTGGACCACTGCTTCTTCCCGGACACGGGTGTGAGCCTGTGCGCGG	386	

Db 1389 TGTGTGTACCGAGGGCTCGGGGGCTTCGTTCACAGAAAGCGTTCCCGGATTTTGT 1448  
Qy 1467 ACAAGAGGATCAAGCTCGCAACAGGCTGGCGAGATCCACGGTGTGCCCGCTGGGGCTGT 1526  
Db 1449 ATAAAGAAATAAGGCGAGGACAGGCTGGCGAGGTGCGAGGTGTTCCACAGGGGGCTGT 1508  
Qy 1527 ATGACGGGCCGCTCCAGAGGTGATGCTGCTGCAAGCGAGGAGTGGCGCTCGGGCC- 1585  
Db 1509 ATGACGGGCCGCTCCAGAGGTGCTGCTGCAAGCGAGGAGTGGCGCCACCATCGGGCCT 1568  
Qy 1586 -----CGCGGCTCTGCGCCAGGCAAGATCTCGTGCCTCTGCTGGCGCAACTACACTAGT 1640  
Db 1569 CAGAGATCGGGCTGCGCGCGGCAAGTGGCCCGCTGCGCGCAACTGACAGT 1628  
Qy 1641 CGGGGTTTCAGCTATCTGGGCTCTCAGGCTGATGACACATCGCCCGACGC 1690  
Db 1629 CGGGGTTTCAGCTCTCGGGCTCTCAGGCTGACGACACAGTGGCCCGCGC 1678

RESULT 13  
LOCUS BX950824 1880 bp mRNA linear VRT 17-FEB-2004  
DEFINITION Gallus gallus finished cDNA, clone CHEST48j16.  
ACCESSION BX950824  
VERSION BX950824.1 GI:42600509  
KEYWORDS  
SOURCE Gallus gallus (chicken)  
ORGANISM Gallus gallus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Archosauria; Aves; Neognathae; Galliformes; Phasianidae;  
Phasianinae; Gallus.  
1 (bases 1 to 1880)  
Boardman,P.E., Bonfield,J.K., Brown,W.R.A., Carder,C., Chalk,S.E.,  
Croning,M.D.R., Davies,R.M., Prancie,M.D., Grafham,D.V.,  
Hubbard,S.J., Humphray,S.J., Hunt,P.J., Maddison,M., McLaren,S.R.,  
Niblett,D., Overton,I.M., Rogers,J., Scott,C.E., Taylor,R.G.,  
Tickle,C. and Wilson,S.A.  
Direct Submission  
Submitted (16-FEB-2004) Sanger Institute, Hinxton, Cambridgeshire,  
CB10 1SA, UK. E-mail enquiries: chickes@bms.umist.ac.uk  
BBSRC/Dundee/Nottingham/Sanger/Sheffield/UMIST Gallus cDNA  
sequencing project.  
This sequence is from the  
BBSRC/Dundee/Nottingham/Sanger/Sheffield/UMIST cDNA collection,  
from a library constructed by Elizabeth Bosch. cDNA was prepared  
from RNA extracted from whole embryo, normalised, and poly  
A-tailed. EcoRI-NotI cut cDNA was then ligated into the vector.  
Vector: pBluescript II KS(+); Site\_1: EcoRI; Site\_2: NotI Host:  
Escherichia coli DH10B.  
Location/Qualifiers  
1..1880  
/organism="Gallus gallus"  
/mol\_type="mRNA"  
/strain="White Leghorn, Hisex"  
/db\_xref="taxon:9031"  
/clone="CHEST48j16"  
/clone\_lib="CSE0CHN04"  
/dev\_stage="stage 10"

FEATURES  
source

ORIGIN

Query Match 63.1%; Score 1066.8; DB 5; Length 1880;  
Best Local Similarity 77.9%; Pred. No. 1e-163;  
Matches 1300; Conservative 0; Mismatches 362; Indels 6; Gaps 1;  
Qy 29 AGATGTCCTTCGCGGCAAGAAAGCATCCCCGGATCAGAGTGACCGCCCTTCGTATC 88  
Db 1 AAGATGTGAAACCGCGGCAAGAGAGCGGTCCCCCGGCGACGAGCGAGCTGCTGATC 60  
Qy 89 AGAGGTGGGAGGATCGTGAATGACGACCTCTTTTACCTGATGTGCACTGGAGAT 148  
Db 61 AAGGGGGGGAATAATCGTCAACAGCAGCAGTCAGTCGTTTTTATGCCGACATTTATGGAAGAT 120  
Qy 149 GGCTTGATAAAACAAATCGGAGAAACCTCATGTCCTCGGGGCGCATCAAGACCATTGAC 208

121 GGCTTGATAAAACAAATCGGAGAAACCTGGCTGTTCCTCCAGCGGGGTGAGACGGTGGAT 180  
209 GCCCAGCGGCTGATGTGCTTCCTGTGTGGGTGAGCTCCACAAAGGCTGAGATGSCCT 268  
181 GCCTACGGGAGAGCTGTGTGTGCGCGGCGGATCGATGTGCAACAGCGGCTGCAATGGCT 240  
269 GTCTCTGGGCATGACACCGGCTGACGATCTTCTGTGAGGGCAACCAAGGCGAGCGCTAGCAGGA 328  
241 GTGATGGGGATGCGCTCTGCTGATGACTTCTACCAAGGCACGAGGCGCGCTGGCGGGA 300  
329 GGAAACCACTGATCTTGTGACCACTGCTTCCTCCGACACCGGCTGTGAGCTGTCTGGCGGCC 388  
301 GGAAACCACTGATTAATGAGCACGCTGTCTGCGCGAGCTGGGATGAGCTGTCTGGCTGCC 360  
389 TAGCAGCAGTGGCGGAGCGGGCGGACAGCGCGGCTGTCTGCGCATCTCTCCCTGACGCTG 448  
361 TAGCAGCAGTGGCGGAGCGGCTGTCTGACGAGGCGCTGTCTGCGCATCTGCGGCTGCACTC 420  
449 GACATCACTCCGATGGCATGAGAGCATCAAGGAGAGCTGTGAGGCGCTCTGTCTCAAGGAGGA 508  
421 GACATCCCGCTGGCATGAGAGCTGTGAGAGCGCTGTGAGGCGCTGTGTAAGGACAAG 480  
509 GGTGTGAATCTCTTCTGTGCTTTCATGCGCATACAGGACCGGTGCGCATGTGCGAGGACAGC 568  
481 GGTGTGAATCTCTTCTGTGCTTTCATGCGCATACAGGACAGCTGCAAGTGCACCGACGCC 540  
569 CAGATGTAGAGATCTTTCAGCATCATCCGGGACCTGCGGCGCTGTGCGCCAGGCTGCACGCT 628  
541 CAGATGTATGAATAATTCGTGATTTATGAGAGCTGTGAGGCGCTGTGCGGCGCTGCACGCT 600  
629 GAGAACGGGGACATCTGTGAGGAGGAGGAGAGCGGTGTGTGAGGCTGTGCGCATCTGCGC 688  
601 GAGAACGGGGACATCTTGTGAGGAGGAGGAGAGGCTGTGTGACATTTGGGATCATTGGG 660  
689 CCCGAGGGCCACTGTCTGAGCCACCCCGAGGAGGTGAGGCTGTGAGGCGGTGTACCGAGCT 748  
661 CCAGAGGGGCGCTGCTGAGCGCGCTGTGAGGAGGTGTAAGAGCAGAGCGGTGTACCGGTGCA 720  
749 GTCCACCATCGCCAAAGCAGGACCAACTGCGCTGTGTGCTACCAAGGTGTAGGCGAGGGG 808  
721 ATTACAAATAGCAAAACAAGCCAACTGCTGCTGCTACCAAAATAATGAGCAAAAGT 780  
809 GCGGCGGAGCGCATCTGCTCAGGCGCAAGCGCAGAGGGGTGTGCTGTGTTGGGGAGCCCATC 868  
781 GCAGCAGATGTGTGCTGTCTCAAGCAAGAGAGAGGCACTGTGTGTGTACGGGGAGCCATC 840  
869 ACCGCGAGCTGGGCAACGAGGTTTCACTACTGTGAGCAAGAACTGGGGCAAGGCTGCA 928  
841 ACTGCCAGCTGGGTGTCTGACGCGCTGCGCATCTGTGAGGCAAGAACTGGGGCAAGGCTGCA 900  
929 GCCTTGTGTGATCAGCCCTGCTCAACCCAGACCCCAAGCGGCGGCGGCGGCGGCGGCGG 988  
901 GCCTTGTGTGATCAGCCCTGCTCAACCCCGCTGAGCGGCGGCGGCGGCGGCGGCGGCGG 960  
989 TTGCTGTCTGAGCGGGGACCTCCAGGTGACAGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1048  
961 CTGCTGTCTGTGGGGACCTGTGAGGTGGCAGGCGAGTGTCTCACTGACGTTTCAACCGGCC 1020  
1049 CAGAAAGGCTGTGGGCAAGGACAACTTTCGGCTGTATCCCGAGGCGGCGGCGGCGGCGGCGG 1108  
1021 CAGAAAGGCTGTGGGCAAGGACAACTTTCACCTCATCCCGAGGCGGCGGCGGCGGCGGCGG 1080  
1109 GAGCGCATGTGATGTCTGGGAGAAATGTGTGGCTCTTGGGAGATGGACGAGATGAG 1168  
1081 GAGCGCATGTGATGTCTGGGAGAAATGTGTGGCTCTTCAAGGAGATGGATGAGAAAGAT 1140  
1169 TTGCTGTGCTGACAGTACAAATGTGTGCAAAATCTTCAATTTTACCCCAAGGAGGGG 1228  
1141 TTGCTGTGCTGACAGGACAAATGTGTGCAAAATCTTCAATTTTACCCCAAGGAGGGG 1200  
1229 CGAGTGGCTGTGGGCTGTGACGCTGTGATGTGGAACCCCAAGGCGGCGGCGGCGGCGGCGG 1288

Db 1201 CGAATCGCGTGGGCTCTGATGACAGACTTGGTCTTGTGGAAACCCAGGGCTACAAAAGTC 1260

Qy 1289 ATCTTCCGACAGACCCACATCTGAAGCTGGAGTACAACTCTCGAGGAGTGGAGTGC 1348

Db 1261 ATCTCGGCAAAAACCCACATTTTGAATGTGGAGTACAACTATTTGAAGGCACGGAGTGT 1320

Qy 1349 CGGGGAGCGCTCGCGTGGTCTAAAGTCAGGGCCGAGTGGCGTGTGAGGACGGGAAGATG 1408

Db 1321 CACGGCGCTCTCTGCTGTGTCATCAGTCAGGGCAAAAGTCTTCTTGAAGACGGGAACCTG 1380

Qy 1409 TTTGTCAACCCGGGGGGGGCGCTTCTGCTCCCTCGGAAACAATCCCGGACTTTTGTCTAC 1468

Db 1381 TTTGTCAACCGAGGCTCGGGGCGCTTCTGTTCCCGAAGAGAGTTCGCGGATTTTGTAT 1440

Qy 1469 AAGAGATCAAGCTCGCAACAGCTGGCGGAGATCCAGGTGTGCCCGTGGCGTGTAT 1528

Db 1441 AAGAAGATAAAGCGAGGAACAGCTGGCGGAGTGCACGGTGTCCCAACGGGGCGTGTAT 1500

Qy 1529 GACGGGCGCTCCACGAGGTGATGGTGCCTGCCAAGCCAGGAGTGGCGCTCCGGCC--- 1585

Db 1501 GACGGGCGGTCCACGATGTCTGCTCAGCAACAAAGCTGTGCCACCATCGCGGCTCC 1560

Qy 1586 ---CGGCGCTCTGCCAGCAAGATCTCGTGCCTCTGTGGCGCAACCTACATCAGTCG 1642

Db 1561 AGGATCGGGCTCGCGCGGCAAGTGCCTGCGCGTGCCTGCGCAACCTGCACCACTG 1620

Qy 1643 GGGTTAGCTATCTGGGTCTAGGCTGATGACCAATCCCGCGAGC 1690

Db 1621 GGGTTAGCTGTCTGGGCTCGAGGCTGACCAACCACTGTCCTCCCGGCGC 1668

## RESULT 14

AF301551  
LOCUS 2741 bp mRNA linear VRT 06-DEC-2003  
DEFINITION Gallus gallus collapsin response mediator protein-3A (CRMP3A) mRNA, complete cds.  
ACCESSION AF301551  
VERSION AF301551.1 GI:33340026  
KEYWORDS Gallus gallus (chicken)  
SOURCE Gallus gallus  
ORGANISM Gallus gallus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae; Gallus.  
REFERENCE 1 (bases 1 to 2741)  
Yuaasa-Kawada, J., Suzuki, R., Kano, F., Ohkawara, T., Murata, M. and Noda, M.  
Axonal morphogenesis controlled by antagonistic roles of two CRMP subtypes in microtubule organization  
Eur. J. Neurosci. 17 (11), 2329-2343 (2003)  
JOURNAL MEDLINE  
22699123  
PUBMED 12814366  
REFERENCE 2 (bases 1 to 2741)  
Yuaasa-Kawada, J. and Noda, M.  
Direct Submission  
Submitted (01-SEP-2000) Molecular Neurobiology, National Institute for Basic Biology, 38 Nishigonaka, Myodaiji-cho, Okazaki, Aichi 444-8585, Japan  
JOURNAL  
FEATURES  
source Location/Qualifiers  
1. .2741  
/organism="Gallus gallus"  
/mol\_type="mRNA"  
/db\_xref="taxon:9031"  
/notes="breed: White leghorn"  
1. .2741  
/gene="CRMP3A"  
67. .2019  
/gene="CRMP3A"  
/codon\_start=1  
/product="collapsin response mediator protein-3A"  
/protein\_id="AAQ14489.1"  
/db\_xref="GI:33340027"  
/translation="MAERSAEPLPASFAGLIGAGSPRQRKGMFCVSDAPDNKTLID

## ORIGIN

Query Match 61.5%; Score 1038.6; DB 5; Length 2741;  
Best Local Similarity 77.9%; Pred. No. 3.6e-159;  
Matches 1267; Conservative 0; Mismatches 354; Indels 6; Gaps 1;

Qy 70 GAGTGACCGCTTCTGATCAGAGGTGGGAGGATCGTGAATGACGACGACGAGTCCCTTTTACGC 129

Db 333 GAGCGAGCAGCTGCTGATCAAGGGGGGAAATCGTCAAGCAGCAGCAGTCTGTTTATGC 392

Qy 130 TGATGTGACGTGGAAAGATGGCTTGATAAAAACAATCGGAGAAACCTCATCGTCCCTGG 189

Db 393 CGACATTTATGTGGAAGATGGGTGATAAAAACAGATTGGAGAGAACTGGCTGTTCCTCAG 452

Qy 190 GGGCATCAAGACCATTTGACGCCACCGCTGATGGTCTTCTTGGTGGCGTTGACCTCCA 249

Db 453 CGGGGTGAGGACGGTGGATGCCCTATGGGAGCTGGTGGTGGCGGCGGATCGATGTGCA 512

Qy 250 CACAGGCTGCGAGATGCTGTCTGGGCGATGACACCGGCTGACGAGCTTCTGTCAAGGCGAC 309

Db 513 CACGGGCTGCGAGTGGCTGTGATGGGATGGCTCTGCTGATGACTTCTTACCAAGGCAC 572

Qy 310 CAAAGCAGCGCTAGCAGGAGGAACACCATGATCTTTGGACCAAGTCTTCCCGACACGGG 369

Db 573 GAGGGCGGCTTGGCGGGAGGAACACCATGATTAATGGACCAAGTCTGCTGCGGAGGTGG 632

Qy 370 TGTGAGCTGCTGGCGGCTTACGAGCAGTGGCGGGAGCGGGGACAGCGCGCTGCTG 429

Db 633 GACGAGCTGCTGGCTGCTTACGAGCAGTGGCGGAGCGGTGTGACAGAGGCGCTGCTG 692

Qy 430 CGACTACTCCCTGCGACGTGGACATCACCCGATGGCATGAGAGCATCAAGAGAGAGCTGGA 489

Db 693 CGACTACGCGCTGCACATCGACATCCCCGCTGGCACAGAGAGCTCGCGGAGGAGCTGGA 752

Qy 490 GGCCTTGGTCAAGAGAGAGGGTGTGAATCTTCTTCTGGTCTTCTATGGCATAAAGACCG 549

Db 753 GCGCTTGGTGAAGGACAGGGTGTGAATCTTCTTCTGGTCTTCTATGGCCTTACAAGGACAG 812

Qy 550 GTGCCAGTGCAGCAGCAGCAGATGTACGAGATCTTTCAGCATCATCCGGGACCTGGGGGC 609

Db 813 ACTGCAGTGCAGCAGCAGCAGATGTGAATATTTCTGCAATATTCGAGACCTGGGGGC 872

Qy 610 CTTGGGCCAGGTGCGACGCTGAGAAACGGGACATCTGTGGAGGAGGAGCAGAAAGCGTGTCT 669

Db 873 CATAGCCCAAGTGCACGCTGAAAACGGGACATCATTTGAGGAGGAGCAGAAAGGCTGCT 932

Qy 670 GGAGCTCGGATCATCTGGGCCCGAGGGCCACGTGTCTGCTGAGCCACCCGAGAGAGTGGAGGC 739

Db 933 GGACATTTGGATCATCTGGGCCAGAGGGGATGTCTTCTGAGCGCGCTTGAAGAGCTTGAAGC 992

Qy 730 TGAGGCGGTGTACCGAGCTGTCCATCGCCAGCAGGAGCAAACTGCCCGCTGTACGTCAC 789

Db 993 AGAGGCGGTCTACCGTGCAATTAACAATAGAAAACAGCCACTGCCCCCTTGTACGTCAC 1052

Qy 790 CAAAGTGTAGCAGCAAGGGGCGGCGCGACGCTATCGCTCAGGGCCAGCGAGAGGGTGGT 849

Db 1053 CAAAATAATGAGCAAAAGTGCACAGATGTGGTAGTCTCAAGCAAAAGAGAAAGGCACGT 1112

Qy 850 CGTGTGTTGGGAGCCCATCACCGCCAGCTTGGGACCGAGCGTTCACATCTACTGGAGCAA 909

Db 1113 GGTGTGAGGAGGCCCATCACTCCAGCCTGGGTGCTGACGGCTCGCACTACTCTGGAGCAA 1172

PFSLGGGRSGSATPSDSSSPAPPPRSPGAATAGMSDGCROSEQLLIGKGIYN  
DDQSFADLYVEDGLIKOIGENLAVPSGVRTVDAYGQLVPGIDVHTRLOWAVMGMA  
RHESRELEALVKDGVNSFLVFMAYKDRLCQTDQMYEIFCIIRDIGALIAQVAHE  
NGDIIEBQKRLLDITGTGPGHVLSPREEVEAAVYRAITTAQNANCPFLYVTKIMSK  
SAADYVQAQRKGTVVYGEPTTASIGASGSHYSNKNMAKAAAFVTPSPSPDPTTFR  
LSLLSCGDLQVAGAHCTFTTAQKAVGKDNFTLIPGTNGIIEERMAIIEWKCVSGEK  
MDENDFVATSTNAAKIENCYPRKGRIVAGSDADLVLMNPRATKVIKSAKTNLVRYN  
IFEGTECHGAPAVKISQKVLVEDGNLFTVEGSGRFPVPRKTFPFYKIKAKRNLAE  
VHGVRGLYDGPVHDVLTSTKAVPTTAAARIAACAKKPAVVPVRNLHQSGFSLSGQA  
DDHVARRTAOKIMAPPGRSNTLS\*

Qy	910	GAATGGGCGGAGGCTGAGGCTTCTGTCACATACCCCTGTGTCAACCCAGACCCACAC	969
Db	1173	GAATGGGCGGAGGCTGAGGCTTCTGTCACATACCCCTGTGTCAACCCAGACCCAC	1232
Qy	970	GGCAGACACCTCACTGCTTCTGTCAGCGGGGAGCTCCAGGTGACGAGCGGCCA	1029
Db	1233	GCCAGAGGCTTCTCTCCCTGCTGTCTGTGGGAGCTGCGAGTGGGAGGAGTGTCTCA	1292
Qy	1030	CTGCACCTTACCACTGCCCAAGAGGCTGTGGGCAAGCAACTTCCGCTGTATCCCGGA	1089
Db	1293	CTGCACCTTACCACTGCCCAAGAGGCTGTGGGCAAGCAACTTCCGCTGTATCCCGGA	1352
Qy	1090	GGGACCAACGGCAATGAGAGGCGCATGTGATGCTGTGGGAGAAATGTGTGGCTCTGG	1149
Db	1353	GGGACCAATGGCAATGAGAGGCGCATGTGATGCTGTGGGAGAAATGTGTGGCTCTGG	1412
Qy	1150	GAAGATGAGAGGAGAAATGAGTTCGTCGGGCTCTGAGCGCTGACGCTGCTCAATCTTCAA	1209
Db	1413	GAAGATGAGAGGAGAAATGAGTTCGTCGGGCTCTGAGCGCTGACGCTGCTCAATCTTCAA	1472
Qy	1210	TTTTTACCAAGGAGGCGAGTGGCTGTGGGCTCTGAGCGCTGACGCTGCTCAATCTTCAA	1269
Db	1473	CTGCTACCCAGGAGGCGGAGTGGCTGTGGGCTCTGAGCGCTGACGCTGCTCAATCTTCAA	1532
Qy	1270	CCCCAAGGCCACCAAGATCATCTCTGCAAGACCCCAAACTCTGAACTGTGAGTCAACAT	1329
Db	1533	CCCCAAGGCCACCAAGATCATCTCTGCAAGACCCCAAACTCTGAACTGTGAGTCAACAT	1592
Qy	1330	CTTCAGGAGTGTGAGTGGCGGAGCGCTGCGCTGCTGATGATGATGATGATGATGATGATG	1389
Db	1593	ATTTGAAGGACCGAGTGTCAAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1652
Qy	1390	GCTGAGGAGCGGAGGATGTTGTGACCCCGGGGGCGGCGCTTCTGCTGCTGCTGCTGCTGCTG	1449
Db	1653	TCTGAGGAGCGGAGGATGTTGTGACCGGAGGCTGCGGGGCTTCTGCTGCTGCTGCTGCTGCTG	1712
Qy	1450	ATTCCCGGACTTGTCTACAGAGGATCAAGTCTGCAAGACCCCAAACTGTGAGTCAACAT	1509
Db	1713	GTTCGCGGATTTGTGTTATTAAGAAATAAAGCGGAGGAAACAGGCTGCGGAGGTGACGG	1772
Qy	1510	TGTGCGGCTGCTGCTGATGAGCGGCGGCTGACGAGGTGATGCTGCTGCTGCTGCTGCTGCTG	1569
Db	1773	TGTGCGGCTGCTGCTGATGAGCGGCGGCTGACGAGGTGATGCTGCTGCTGCTGCTGCTGCTG	1832
Qy	1570	GAGTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1623
Db	1833	GCCACCATCGGCGCTCCAGGATCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1892
Qy	1624	GCGCAACTACATCAGTGGGCTTCAAGCTTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1683
Db	1893	GCGCAACTGACCACTGCGGCTTCAAGCTTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1952
Qy	1684	CCGAGCG 1690	
Db	1953	CCGAGCG 1959	
RESULT 15			
BTU83278			
LOCUS			
DEFINITION Bos taurus neural specific protein CRMP-2 gene, complete cds.			
ACCESSION U83278			
VERSION U83278.1 GI:1916226			
KEYWORDS			
SOURCE Bos taurus (cow)			
ORGANISM			
Bos taurus			
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;			
Bovinae; Bos.			
1 (bases 1 to 3884)			
REFERENCE			
AUTHORS			
TITLE			
Isolation and characterization of a neural specific protein (Nep60)			
gene homologous to C. elegans unc-33 controlling axonal outgrowth			

JOURNAL  
REFERENCE  
2 (bases 1 to 3884)  
Kamata, T. K.  
Direct Submission  
Submitted (27-DEC-1996) Lab. of Biochemical Physiology, National  
Cancer Inst. FCRDC, P.O. Box B, Bldg 567, Frederick, MD 21703, USA

and guidance  
Brain Res. Mol. Brain Res. (1997) In press  
Kamata, T. K.  
Direct Submission  
Submitted (27-DEC-1996) Lab. of Biochemical Physiology, National  
Cancer Inst. FCRDC, P.O. Box B, Bldg 567, Frederick, MD 21703, USA

FEATURES  
source

1. 3884

/organism="Bos taurus"

/mol\_type="genomic DNA"

/db\_xref="taxon:9913"

/notes="map location 8p21 in human; mouse 14"

177. 1895

/notes="unc-33 homolog"

/codon\_start=1

/product="neural specific protein CRMP-2"

/protein\_id="AA80618.1"

/db\_xref="GI:1916227"

/translation="MSYQKKNIPIRITSLRLIKGGKIVNDQSFYADIYMBDLIKQ  
IGENLVPSGKTIHAHSMVPGIDVHTRFQPMQGMSTADDFQGTKAALAGGTT  
MIDHVPSPGTSLLAAPQWRWADSKCCDYSLHVDITMHHKGVQBEAEALVKDHG  
VNSPLVMAFKDQFQTLDSQIYEVLSVIRDIGAIAQVHAENGDIIEEQRIIDLGIT  
GPEGHVLSRPEBEAARVSIITANGTNCPLVITTKVMSAAVIAQARKKGTVVVG  
EPITASLGDTGSHYNSKNWAKAAAFVTSPLSPDTPDPLNSLLSGDQLQVTSVHC  
TFNPAKRTAVGSDADLVWDPSVKTISAKTHNSLEYNIPEGMCRGSPVAVISQV  
NLVPRGTLVHTGSGRYIPKRPDPVVKIKARSLAELRGVPRGLVDGVPCEVSV  
TKPTVTPASAKTSAPKQAPVPRNLHQSGFSLGAIQIDNIPRTTQRIIVAPGGRA  
NITSIG"

ORIGIN

Query Match 58.9%; Score 995.8; DB 4; Length 3884;  
Best Local Similarity 74.7%; Pred. No. 31e-152;  
Matches 1249; Conservative 0; Mismatches 422; Indels 0; Gaps 0;

Qy 20 CCAGAGGAGGATGCTCTTCCAGGGCAAGAAAGCATCCCGGATTCACAGGTGACCGC 79  
Db 165 CCAGAGGAGGATGCTCTTATCAGGGAGAGAAAATATTCAGGCATCAGAGTGATCGT 224  
Qy 80 CTTCTCATGAGGTGGGAGGATCGTGAATGACGACGACGCTCTTTTACGCTGATGTGAC 139  
Db 225 CTTCTCATCAAGAGGATGATTAATGATGACGACGCTCTTTATGACGACATATAC 284  
Qy 140 GTGCAAGATGCTGTTGATAAAACAAATCGGAGAAACCTCATCGTCCCTGGGGCATCAAG 199  
Db 285 ATGCAAGATGCTGTTGATCAAGCAATAGGAGAAAACCTGATTTGCGGAGGGGTGAG 344  
Qy 200 ACCATTGACGCCACCGCTGATGGTCTCTTCTGGTGGCGTTGACGCTCCACACAAGCTG 259  
Db 345 ACCATTGAGGCCATTCACAGATGGTGATCCCTGGAGGATTCAGCTCCATACCCGCTTC 404  
Qy 260 CAGATGCTGCTCTGGGCATGACACCGGCTGACGACTTCTGTGAGGGCACCAGGGACGG 319  
Db 405 CAGATGCTGATCAGGGGATGACCTCTGCTGATGACTTCTTTTCAAGGCACCAAGGACGC 464  
Qy 320 CTAGCAGGAGGAAACACCATGATCTTGGACACGCTCTTCCCGACACGCGGTGTGAGCTG 379  
Db 465 CTGGCTGAGGAAACACCATGATCTTGGACACGCTGTTCTGAGCTGGGACAGGCTG 524  
Qy 380 CTGGCGGCTTACAGCAGTGGCGGAGCGGCGGACAGCGCGCTGCTGCTGCTGCTGCTGCTG 439  
Db 525 CTGGCGGCTTTCACAGTGGCGGAAATGGCTGACAGCAAGTCTCTGCTGCTGCTGCTGCTG 584  
Qy 440 CTGCACTGGAACATCAACCGATGCGATGAGAGCATCAAGGAGGCTGAGGGCTCTGCTGCT 499  
Db 585 CTGCACTGGAACATCACTGAATGGCAAGGGGCTTCAGGAGAGATGAGGGCTCTGCTGCTG 644  
Qy 500 AAGGAGAGGCTGTAACCTCTTCTGCTGCTTCTCATGCGATACAGGACCGGTGCTGCTGCTG 559  
Db 645 AAGGAGAGGCTGTAACCTCTTCTGCTGCTTCTCATGCGATACAGGACCGGTGCTGCTGCTG 704  
Qy 560 AGCGACAGCCAGATGTACAGAGATCTTTCAGCATCATCCGGACCTGGGGGCTTGGGCCAG 619

```
Db 705 ACAGATTCCAGATCTATGAAGTCTCTGAGCGTATCCGGGATATGGAGGATGGCCAG 764
Qy 620 GTGCAAGCTGAGAACGGGGGACATCTGTGGAGAGGAGAGAGCGGTTGTGGAGCTCGGC 679
Db 765 GTCCACGCGGAAACGGCGACATCATCGCAGAGAGCAGCAGAGGATCTCTGGATCTGGGC 824
Qy 680 ATCACTGGCCCCGAGGGCCAGCGTCTCAGCACCCGAGGAGGTGGAGGCTGAGGGCGGTG 739
Db 825 ATCAAGGGCCCTGAGGGACATGTGTGAGCGCGCTGAGGAGGTGAGGGCTGAAGCCGTG 884
Qy 740 TACCGAGCTGTACCATCGCAAGCAGGCAAACTGCCCGCTGTACGCTCACCAAGGTGATG 799
Db 885 AACGGTTCCATCAACCATCGCAACCAAGACCACTGCGCTCTGTACATCACCAAGGTGATG 944
Qy 800 AGCAAGGGGGCGCGCAGCGCATCGCTCAGGCCAAGCGCAGAGGGGTGTGCTGTTTGGG 859
Db 945 AGCAAGAGCGCAGCGGAAGTTATCGCCCCAGCGCCCGGAAAGGGAAACCGTGTGTACGGC 1004
Qy 860 GAGCCCATCACCGCCAGCCTGGGCAACCGAGCGTTTCACTACTGGAGCAAGAACTGGGC 919
Db 1005 GAGCCCATCACCGCCAGCCTGGGCAACCGAGCGGTCTCTACTACTGGAGCAAGAACTGGGC 1064
Qy 920 AAGGCTGACGCTTGTGTACATCACCCCTGTCAACCCAGACCCCAACACGCGCAGACCAC 979
Db 1065 AAGGCTGCTGCTTGTGTACCTCTCCCGCCCTGAGCCCGGACCCCAACCACTCCAGACTTT 1124
Qy 980 CTCACTGTGTGTGTCCAGCGGGACCTCAGGTGACAGGCGAGCGGCCACTGACCTTC 1039
Db 1125 CTGAACTCTTGTGTGTGTGGAGATCTCAGGTGACCGGCGAGTGTCTACTGTACCTTC 1184
Qy 1040 ACCACTGCCCAGAGGCTGTGGCAGAGGACAACTTCGGCTGTATCCCGGAGGGCACAAC 1099
Db 1185 AACACGGCCAGAGGCGGTGGAAAGACAACCTTCACTCCCTGTATCCCTGAGGGGACCAAC 1244
Qy 1100 GGCATTGAGGAGCGCATGTGATGTGGGAGAAATGTGTGGCTCTGGGAAAGATGGAC 1159
Db 1245 GGCACGAGGAGCGGATGTCTGTCTATCTGGGATAGAGCTGTGTCTACTGGGAGATGGAC 1304
Qy 1160 GAGATGAGTGTGTGCGGTGACAGTACAAATGTGTGCCAAATCTTCAATTTTACCCA 1219
Db 1305 GAGAACCAAGTTGTGGCGGTGACAGCACCAATGTGTGCCAAAGTCTTCAACCTGTACCCC 1364
Qy 1220 AGGAGGGGAGTGTGGCTGTGGCTCTGACGCTGATCTGTGATATGGAACCCCAAGGCC 1279
Db 1365 CGGAAGGGCGCATTGTGTGGGTGTGATGTGTGATCTGTGGACCCCGCAGC 1424
Qy 1280 ACCAAGATCATCTTGCAAGACCCCAATCTGAACGTGGAGTACAAATCTTTCGAGGGA 1339
Db 1425 GTGAGACCATCTCTGCCAAGACCCCAACAGCTCTCTTCAATACAACTTTCGAAGGC 1484
Qy 1340 GTGAGTGGCGGAGCGGCTGTGGTGTGATGATGAGTCAAGGCGCGAGTGGCGCTGGAGGAC 1399
Db 1485 ATGAGTGGCGGAGGCTCCCACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1544
Qy 1400 GGGAGAGTGTGTGTACCCCGGGGGCGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1459
Db 1545 GGTACCCCTTCAATGTCAAGAGGCTCTGGGCGCTACATCCCGGAGAGCCCTTCCCTGAC 1604
Qy 1460 TTTGTCTCAAGAGGATCAAGCTCGCAACAGGCTGGCGGAGATCCACGGTGTGCCCGCT 1519
Db 1605 TTCGTTTCAAGCGTATCAAGGCAAGGACAGGCTGGCGGAGCTGAGAGGGGTTCCTCGT 1664
Qy 1520 GGGCTGTATGACGGGCGGCTGACAGGTTATGTTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1579
Db 1665 GGTCTGTACGACGACCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1724
Qy 1580 CCGGCGCGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1639
Db 1725 TCCTGGGCTAAGAGGTCTCGGCGCAAGCAGAGGCGCCCACTGTGTGTGTGTGTGTGTGTGT 1784
Qy 1640 TCGGGGTTCAGGCTTATCTGGGTCTCAGGCTGATGACCAATCGCCCGACGC 1690
```

Db 1785 TCTGGATTCAAGTGTGTTCGGTGCAGATTTGATGACAAATCCTCGCGC 1835

Search completed: September 24, 2005, 07:02:31  
Job time : 24238 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on:

September 23, 2005, 23:03:07 ; Search time 3115 Seconds  
(without alignments)  
3211.673 Million cell updates/sec

Title:

US-09-367-496C-7

Perfect score:

1690

Sequence:

1 gccgcctaccagagacc.....atgaccacatgcccgacgc 1690

Scoring table:

IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched:

4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters:

8780412

Minimum DB seq length:

0

Maximum DB seq length:

2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

N\_Geneseq\_16Dec04:\*  
1: Geneseqn1980s:\*  
2: Geneseqn1990s:\*  
3: Geneseqn2000s:\*  
4: Geneseqn2001as:\*  
5: Geneseqn2001bs:\*  
6: Geneseqn2002as:\*  
7: Geneseqn2002bs:\*  
8: Geneseqn2003as:\*  
9: Geneseqn2003bs:\*  
10: Geneseqn2003cs:\*  
11: Geneseqn2003ds:\*  
12: Geneseqn2004as:\*  
13: Geneseqn2004bs:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1690	100.0	1690	6	ABK91191 Human cdn
2	1688.4	99.9	1690	2	AAV60818 Human par
3	1688.4	99.9	2741	12	ADQ23110 Human sof
4	1686.8	99.8	2699	12	ADQ18667 Human sof
5	1318.4	78.0	2352	9	AD57506 Human enz
6	1303.6	77.1	1920	2	AAV60817 Mouse ULI
7	1067.2	63.1	2470	12	ADK70708 Collapsin
8	1038.6	61.5	2741	12	ADK70695 Chicken C
9	980.2	58.0	2947	12	ADJ84229 Malayan b
10	970.6	57.4	1817	2	AAV60815 Mouse ULI
11	965.4	57.1	1829	6	ABK91188 Human cdn
12	963.8	57.0	4459	9	ACF25352 Human dii
13	963.8	57.0	4459	10	ADC66333 Human col
14	963.8	57.0	4459	13	ADR98995 Dihydroxy
15	963.8	57.0	5421	6	ABL68565 Kidney ca
16	963.8	57.0	5421	6	ABL6868 Lung canc
17	963.8	57.0	5421	10	ADC66325 Human col
18	963.8	57.0	5421	10	ADH28910 Human chr
19	963.8	57.0	5421	13	ADR52683 Drug ther
20	963.8	57.0	5421	13	ADT93859 Non-small

21	929.6	55.0	2393	12	ADK70706	Adk70706 Collapsin
22	902.4	53.4	2488	12	ADK70693	Adk70693 Chicken C
23	879.6	52.0	2202	12	ADK70710	Adk70710 Collapsin
24	878.2	52.0	2297	2	AAV60816	Av60816 Mouse ULI
25	876.8	51.9	2869	10	ADJ56545	Adj56545 Murine cd
26	876.8	51.9	2928	11	ADN95177	Adn95177 Human BEC
27	876.8	51.9	2928	13	ACN38089	Acn38089 Tumour-as
28	875.2	51.8	2797	13	ADR07554	Adr07554 Full leng
29	875.2	51.8	2842	6	ABK91190	Abk91190 Human cdn
30	875.2	51.8	2842	6	ABK39752	Abk39752 cDNA enco
31	875.2	51.8	2842	8	ACA12081	Aca12081 Human lun
32	875.2	51.8	2842	8	ACA03267	Aca03267 Lung canc
33	875.2	51.8	2842	10	ADH47319	Adh47319 Human lun
34	875.2	51.8	2842	13	ADJ21238	Adj21238 Human lun
35	873.4	51.7	1719	8	ACC48980	Acc48980 Human col
36	868.6	51.4	2813	10	ADB59112	Adb59112 Toxicity-
37	850.6	50.3	5046	6	ABL61765	Ab161765 Colon ade
38	850.6	50.3	5047	5	AAH81784	Aah81784 Human dif
39	850.6	50.3	5047	6	ABL65185	Ab165185 Lung canc
40	850.6	50.3	5047	11	ADP65042	Adp65042 Human dii
41	850.6	50.3	5047	11	ADP65567	Adp65567 Human mRN
42	850.6	50.3	5047	12	ADN03726	Adn03726 Antipeori
43	850.6	50.3	5200	10	ADJ56540	Adj56540 Chicken c
44	850.6	50.3	5250	4	AAL26651	Aal26651 Human bre
45	850.6	50.3	5578	11	ACN88655	Acn88655 Breast ca

#### ALIGNMENTS

RESULT 1  
ABK91191  
ID ABK91191 standard; cDNA; 1690 BP.

AC ABK91191;

DT 05-NOV-2002 (first entry)

XX Human cDNA encoding partial Ulip4/CRMP3 protein.

XX Human; ss; gene; Ulip4; CRMP3 collapsin response mediator protein;  
KW Unc-33-like protein; neurodegenerative disease; Alzheimer's disease;  
KW paraneoplastic neurodegenerative disease; PND; myelination;  
KW demyelination; remyelination; myelin disorder; multiple sclerosis;  
KW autoimmune neurodegenerative disorder; HTLV-1 associated myelopathy;  
KW human T lymphocyte virus 1.

XX Homo sapiens.

Key Location/Qualifiers  
CD5 32..1689  
/\*tag= a  
/partial

FT /product= "Amino acids 1-553 of Ulip4/CRMP3"  
FT /note= "This sequence is stated to encode Ulip4/CRMP3  
FT protein appearing as ABG32231, but only encodes amino  
FT acids 1-553"  
FT /transl\_except= (pos:197..199,aa:His)

US2002119944-A1.

29-AUG-2002.

PF 09-NOV-2001; 2001US-00986632.

XX 09-NOV-2000; 2000US-0246751P.

XX (AGUE/) AGUIERA M.

PA (BELI/) BELIN M.

PA (CHAR/) CHARRIER E.

PA (HONO/) HONORAT J.

PA (RICA/) RICHARD D.

PA (ROGE/) ROGEMOND V.

XX Aguera M, Belin M, Charrier E, Honorat J, Ricard D, Rogemond V;  
XX MPI; 2002-627172/67.  
XX Prevention or treatment of myelin disorders, such as multiple sclerosis,  
PT by administering an agent selected from a Ulp1/CRMP protein, a nucleic  
PT acid coding for the protein, or an antibody directed against protein.  
XX  
XX Disclosure; Page 25-26; 44pp; English.  
XX The invention relates to a new method for prevention or treatment of  
CC myelin disorders, comprises administering to a patient an effective  
CC amount of an agent selected from a Ulp1 (Unc-33-like protein)/CRMP  
CC (collapsin response mediator protein) protein, a nucleic acid coding for  
CC Ulp1/CRMP, an antisense sequence capable of specifically hybridizing with  
CC the nucleic acid, an antibody directed against Ulp1/CRMP, or an aptamer  
CC capable of binding Ulp1/CRMP, and a pharmacologically acceptable carrier.  
CC Also included are methods of diagnosing a myelin disorder in a subject,  
CC identifying agents useful for the prevention or treatment of myelin  
CC disorders, using the Ulp1/CRMP proteins/nucleic acids, agents capable of  
CC modulating the function or expression of the proteins (increasing or  
CC decreasing), and a method for identifying an endogenous agent as a  
CC therapeutic target for the prevention or the treatment of myelin  
CC disorders. The agents are useful for preventing or treating a myelin  
CC disorder such as multiple sclerosis or HTLV-1 (human T lymphocyte virus  
CC 1) associated myelopathy and neurodegenerative diseases, Alzheimer's  
CC disease, paraneoplastic neurodegenerative diseases (PND), autoimmune  
CC neurodegenerative disorder. Ulp1/CRMP proteins are involved in  
CC the processes of myelination, demyelination and remyelination. Antibodies  
CC to a Ulp1/CRMP protein are useful for diagnosing a myelin disorder. The  
CC present sequence encodes a partial Ulp1/CRMP3 protein  
XX  
SQ Sequence 1690 BP; 360 A; 505 C; 539 G; 286 T; 0 U; 0 Other;  
Query Match 100.0%; Score 1690; DB 6; Length 1690;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1690; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 GCCGCCCTACAGAGACCCCGAGGACAGATGTCCTTCCAGGGCAAGAAAGCATCC 60  
DB 1 GCCGCCCTACAGAGACCCCGAGGACAGATGTCCTTCCAGGGCAAGAAAGCATCC 60  
QY 61 CCGGATCAGAGTGACCGCTTCTGATCAGAGTGGGAGGATCGTGAATCAGCACCATC 120  
DB 61 CCGGATCAGAGTGACCGCTTCTGATCAGAGTGGGAGGATCGTGAATCAGCACCATC 120  
QY 121 CTTTATCGCTGATGTCACGTGGAAGATGGCTTGATATAAAACAAATCGAGAAACCTCAT 180  
DB 121 CTTTATCGCTGATGTCACGTGGAAGATGGCTTGATATAAAACAAATCGAGAAACCTCAT 180  
QY 181 CGTCCCTGGGGGCATCAAGACCAATTGACGCCACCGCTGATGCTCTTCTGGTGGCGT 240  
DB 181 CGTCCCTGGGGGCATCAAGACCAATTGACGCCACCGCTGATGCTCTTCTGGTGGCGT 240  
QY 241 TGAGCTCCACAAAGGCTGAGATGCTCTGCTGGGATGACACCGCTGACGACTTCTG 300  
DB 241 TGAGCTCCACAAAGGCTGAGATGCTCTGCTGGGATGACACCGCTGACGACTTCTG 300  
QY 301 TCAGGGCACCAGGACGCTAGCAGAGGAAACCAACATGATCTTGGACCAAGCTTCTCC 360  
DB 301 TCAGGGCACCAGGACGCTAGCAGAGGAAACCAACATGATCTTGGACCAAGCTTCTCC 360  
QY 361 CGACAGGGGTGAGCTGCTGGCGCTTACGAGCATGCGGAGCGGCGGACAGCGC 420  
DB 361 CGACAGGGGTGAGCTGCTGGCGCTTACGAGCATGCGGAGCGGCGGACAGCGC 420  
QY 421 GGCTGTGCTGACTACTCTCTGCTGAGCATCACCGGATGGCATGAGGATCAAGGA 480  
DB 421 GGCTGTGCTGACTACTCTCTGCTGAGCATCACCGGATGGCATGAGGATCAAGGA 480  
QY 481 GGAGCTGGAGGCCCTGGTCAAGGAGAGAGGTGTGAATCTCTCTGCTTCTCATGGCATA 540  
DB 481 GGAGCTGGAGGCCCTGGTCAAGGAGAGAGGTGTGAATCTCTCTGCTTCTCATGGCATA 540

DB 481 CGAGCTGGAGGCCCTGGTCAAGGAGAGAGGTGTGAATCTCTCTGCTTCTCATGGCATA 540  
QY 541 CAAAGCACCGTGGCAGTGCAGCGACAGCCAGAGATGACGAGATCTTTCAGCATCATCCGGGA 600  
DB 541 CAAAGCACCGTGGCAGTGCAGCGACAGCCAGAGATGACGAGATCTTTCAGCATCATCCGGGA 600  
QY 601 CTTGGGGGCTTTGGCCCGAGGTGCAGCTCAGAAACGGGGACATCTGTGGAGAGAGAGCAGAA 660  
DB 601 CTTGGGGGCTTTGGCCCGAGGTGCAGCTCAGAAACGGGGACATCTGTGGAGAGAGAGCAGAA 660  
QY 661 GCGGTGTCTGGAGCTCGGCATCACTGGCCCCGAGGGCCACGTCCTCAGCCACCCCGAGGA 720  
DB 661 GCGGTGTCTGGAGCTCGGCATCACTGGCCCCGAGGGCCACGTCCTCAGCCACCCCGAGGA 720  
QY 721 GGTGGAGGCTGAGGCGGTGTACCGAGCTGTACCATTCGCAAGCAGGCAAACTGCCCGCT 780  
DB 721 GGTGGAGGCTGAGGCGGTGTACCGAGCTGTACCATTCGCAAGCAGGCAAACTGCCCGCT 780  
QY 781 GTACGTCAACCAAGGTGATGAGCAAGGGGGCGGCGGCGCATCGCTCAGGCCCAAGCGCAG 840  
DB 781 GTACGTCAACCAAGGTGATGAGCAAGGGGGCGGCGGCGCATCGCTCAGGCCCAAGCGCAG 840  
QY 841 AGGGGTGTCTGTGTTTGGGGAGCCCATCACCGCAGCTTGGGCACGACGGTTTCACTA 900  
DB 841 AGGGGTGTCTGTGTTTGGGGAGCCCATCACCGCAGCTTGGGCACGACGGTTTCACTA 900  
QY 901 CTGGAGCAAGAACTGGGCCAAGGCTGCAGCTTGTGCATCATCACCCCTTGTCAACCCAGA 960  
DB 901 CTGGAGCAAGAACTGGGCCAAGGCTGCAGCTTGTGCATCATCACCCCTTGTCAACCCAGA 960  
QY 961 CCCCACACGGCAGACCACTCCTCCTGCTTGTCTCAGCGGGGACCTCCAGGTGACAGG 1020  
DB 961 CCCCACACGGCAGACCACTCCTCCTGCTTGTCTCAGCGGGGACCTCCAGGTGACAGG 1020  
QY 1021 CAGCGCCCATGCACTTCACTCCACAGAGGCTGTGGGCAAGGACAACTTCGCGCT 1080  
DB 1021 CAGCGCCCATGCACTTCACTCCACAGAGGCTGTGGGCAAGGACAACTTCGCGCT 1080  
QY 1081 GATCCCGAGGGGACCAAGGCAATGAGAGGCGCATGTGATGGTCTGGGAGAAATGTGT 1140  
DB 1081 GATCCCGAGGGGACCAAGGCAATGAGAGGCGCATGTGATGGTCTGGGAGAAATGTGT 1140  
QY 1141 GGCTCTGGAGAGATGCAAGAAATGAGTTGTGCGGGTGACCAAGTACAAATCTGCGCA 1200  
DB 1141 GGCTCTGGAGAGATGCAAGAAATGAGTTGTGCGGGTGACCAAGTACAAATCTGCGCA 1200  
QY 1201 AATCTTCAATTTTACCCAAAGGAGGGGAGTGGCTGTGGGCTTGAAGCTGACCTGCT 1260  
DB 1201 AATCTTCAATTTTACCCAAAGGAGGGGAGTGGCTGTGGGCTTGAAGCTGACCTGCT 1260  
QY 1261 CATATGGAACCCCAAGGCCACCAAGATCATCTCTGCAAGACCAACATCTGAACGTGA 1320  
DB 1261 CATATGGAACCCCAAGGCCACCAAGATCATCTCTGCAAGACCAACATCTGAACGTGA 1320  
QY 1321 GTACAACTCTTCGAGGAGTGGAGTGGCGGGAGCGCTTCCGCTGATTAAGTCAAGG 1380  
DB 1321 GTACAACTCTTCGAGGAGTGGAGTGGCGGGAGCGCTTCCGCTGATTAAGTCAAGG 1380  
QY 1381 CCGAGTGGGCTGCGAGGACGGGAAGATGTTTGTACCCCGGGGGCGGGCGCTTCTGCTCC 1440  
DB 1381 CCGAGTGGGCTGCGAGGACGGGAAGATGTTTGTACCCCGGGGGCGGGCGCTTCTGCTCC 1440  
QY 1441 TCGGAAAACTTCCCGGACTTTGTCTACAAAGGATCAAAAGCTTCGCAACAGGCTGGCGGA 1500  
DB 1441 TCGGAAAACTTCCCGGACTTTGTCTACAAAGGATCAAAAGCTTCGCAACAGGCTGGCGGA 1500  
QY 1501 GATCCACGGTGTGCCCGCTGTGATGACGGGCGCTTCACAGAGTGTATGTGTCTGC 1560  
DB 1501 GATCCACGGTGTGCCCGCTGTGATGACGGGCGCTTCACAGAGTGTATGTGTCTGC 1560  
QY 1561 CAAAGCCAGGAGTGGCGCTCCGGCCCGCGCTCTCCAGGCAAGATCTTCGCTGCTCC 1620  
DB 1561 CAAAGCCAGGAGTGGCGCTCCGGCCCGCGCTCTCCAGGCAAGATCTTCGCTGCTCC 1620





```
Qy 781 GTACGTCAACAGGTGATGACGAAGGGGGGGCGAGCCATCGCTCAGGCGCAAGCGCAG 840
Db 898 GTACGTCAACAGGTGATGACGAAGGGGGGGCGAGCCATCGCTCAGGCGCAAGCGCAG 957
Qy 841 AGGGTGGTGTGTTTGGGAGGCCCATACCGCCAGCGCTGGGCAACCGCGTTCACTA 900
Db 958 AGGGTGGTGTGTTTGGGAGGCCCATACCGCCAGCGCTGGGCAACCGCGTTCACTA 1017
Qy 901 CTGAGCAAGAACTGGGCGCAAGGCTGACAGCTTCTGTCACATCACCCCTGTCAACCCAG 960
Db 1018 CTGAGCAAGAACTGGGCGCAAGGCGCGAGCTTCTGTCACATCACCCCTGTCAACCCAG 1077
Qy 961 CCCACACCGGCAAGCACTCACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1020
Db 1078 CCCACACCGGCAAGCACTCACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1137
Qy 1021 CAGGCGCCACTGCACTTCTCACTCTGCGGCGGAGGCTGTGGGCAAGCAACTTCGGGCT 1080
Db 1138 CAGGCGCCACTGCACTTCTCACTCTGCGGCGGAGGCTGTGGGCAAGCAACTTCGGGCT 1197
Qy 1081 GATCCCCGAGGCGCAACCGCATTTGAGGAGCGCATGTCATGCTCTGGGAGAAATGTGT 1140
Db 1198 GATCCCCGAGGCGCAACCGCATTTGAGGAGCGCATGTCATGCTCTGGGAGAAATGTGT 1257
Qy 1141 GGCTCTTGGGAAGATGGAAGAAATGATGTTCTGTCGGGCTGACCAAGTAAATGCTGCCAA 1200
Db 1258 GGCTCTTGGGAAGATGGAAGAAATGATGTTCTGTCGGGCTGACCAAGTAAATGCTGCCAA 1317
Qy 1201 AATCTTCAATTTTACCAGGAAGGGCGAGTGCTGTGGGCTCTGACGCTGACCTGGT 1260
Db 1318 AATCTTCAATTTTACCAGGAAGGGCGAGTGCTGTGGGCTCTGACGCTGACCTGGT 1377
Qy 1261 CATATGNAACCCCAAGGCCACCAAGATCATCTCTGCCAAGCCACCAATCTGAACTGGA 1320
Db 1378 CATATGNAACCCCAAGGCCACCAAGATCATCTCTGCCAAGCCACCAATCTGAACTGGA 1437
Qy 1321 GTACAACATCTTCAGGGAGTGAGTGCCGGGAGCGCTTCCGCTGTCTATAAGTCAGGG 1380
Db 1438 GTACAACATCTTCAGGGAGTGAGTGCCGGGAGCGCTTCCGCTGTCTATAAGTCAGGG 1497
Qy 1381 CCAGTGGCGCTGAGGAGCGGGAAGATGTTTGTGTCACCCCGGGGGCGGCTTCTGCTCC 1440
Db 1498 CCAGTGGCGCTGAGGAGCGGGAAGATGTTTGTGTCACCCCGGGGGCGGCTTCTGCTCC 1557
Qy 1441 TCGGAAAAATTCGCGACTTTGTCTACAAGAGGATCAAGCTCGCAACAGGCTGGCGGA 1500
Db 1558 TCGGAAAAATTCGCGACTTTGTCTACAAGAGGATCAAGCTCGCAACAGGCTGGCGGA 1617
Qy 1501 GATCCACGGTGTGCCCGCTGGGCTGTATGACGGGCGCGCTCCACGAGGTGATGCTGCTGC 1560
Db 1618 GATCCACGGTGTGCCCGCTGGGCTGTATGACGGGCGCGCTCCACGAGGTGATGCTGCTGC 1677
Qy 1561 CAAGCCAGGAGTGGGCTCCGGCGCGCGCTCTGCGCCAGGCAAGATCTCGGTGCTTCC 1620
Db 1678 CAAGCCAGGAGTGGGCTCCGGCGCGCGCTCTGCGCCAGGCAAGATCTCGGTGCTTCC 1737
Qy 1621 TGTGCGCAACTACATCAGTCTGGGGTTTCAAGCTTATCTGGGTCTCAGGCTGATGACACAT 1680
Db 1738 TGTGCGCAACTACATCAGTCTGGGGTTTCAAGCTTATCTGGGTCTCAGGCTGATGACACAT 1797
Qy 1681 CGCCCGACGC 1690
Db 1798 CGCCCGACGC 1807
```

RESULT 4  
ADQ18657  
ID ADQ18667 standard; DNA; 2699 BP.  
XX  
AC ADQ18667;  
XX  
DT 26-AUG-2004 (first entry)

```
XX Human soft tissue sarcoma-upregulated DNA - SEQ ID 1486.  
XX soft tissue sarcoma; cytostatic; gene therapy; vaccine; screening; human;  
KW db.  
XX Homo sapiens.  
XX WO2004048938-A2.  
XX 10-JUN-2004.  
XX 26-NOV-2003; 2003WO-US038193.  
XX 26-NOV-2002; 2002US-0429739P.  
XX (PROT-) PROTEIN DESIGN LABS INC.  
XX Aziz N, Ginsburg WM, Zlotnik A;  
XX WPI; 2004-441208/41.  
XX Early detection of soft tissue sarcoma comprises determining expression  
XX of a gene in a first soft tissue sample and a normal soft tissue sample  
XX and comparing the gene expression, also useful in treating soft tissue  
XX sarcoma.  
XX Example 2; SEQ ID NO 1486; 210pp; English.  
XX The invention relates to a novel method for detecting soft tissue sarcoma  
XX which comprises obtaining a first soft tissue sample from an individual,  
XX and a normal soft tissue sample from the same or different individual,  
XX determining the expression of a gene in both samples and comparing the  
XX expression of the gene in both soft tissue samples, where a higher level  
XX of protein expression in the first soft tissue sample indicates the  
XX presence of soft tissue sarcoma. The method of the invention has  
XX cytotatic applications and may be useful for detecting soft tissue  
XX sarcoma, possibly via gene therapy or vaccine production. The nucleic  
XX acid sequences may be useful in diagnostic and screening applications.  
XX The current sequence is that of a human soft tissue sarcoma-upregulated  
XX DNA of the invention. The current sequence is not shown within the  
XX specification per se but was submitted in CD format by the inventor.  
SQ Sequence 2699 BP; 535 A; 864 C; 830 G; 470 T; 0 U; 0 Other;
```

```
Query Match 99.8%; Score 1686.8; DB 12; Length 2699;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 1688; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
Qy 1 GCGCGCCCTTACAGAGAGACCCCGAGGAGCAGGATGCTCTCCAGGCGCAAGAAAGCATCCC 60  
Db 118 GCGCGCCCTTACAGAGAGACCCCGAGGAGCAGGATGCTCTCCAGGCGCAAGAAAGCATCCC 177  
Qy 61 CCGGATCAAGAGTACCGCTTCTGATCAGAGGTGGAGGATCGTGAATGACGACCATC 120  
Db 178 CCGGATCAAGAGTACCGCTTCTGATCAGAGGTGGAGGATCGTGAATGACGACCATC 237  
Qy 121 CTTTTCAGCTGATGTCAGCTGAGAGATGCTTGTATAAACAAATCGGAGAAAACCTCAT 180  
Db 238 CTTTTCAGCTGATGTCAGCTGAGAGATGCTTGTATAAACAAATCGGAGAAAACCTCAT 297  
Qy 181 CGTCCCTGGGGGCGATCAAGACCATTCAGCCCGCCATGATGGTCTCTTCTGGTGGCGT 240  
Db 298 CGTCCCTGGGGGCGATCAAGACCATTCAGCCCGCCATGATGGTCTCTTCTGGTGGCGT 357  
Qy 241 TGAAGTCCACAAAGCTCAGATGCTCTCTGGGCGATGACACCGGCTGACGACTTCTG 300  
Db 358 TGAAGTCCACAAAGCTCAGATGCTCTCTGGGCGATGACACCGGCTGACGACTTCTG 417  
Qy 301 TCAGGGCACAAGGCGAGCGCTAGCAGGAGAACCATGATCTTGGACCAAGCTCTTCCC 360  
Db 418 TCAGGGCACAAGGCGAGCGCTAGCAGGAGAACCATGATCTTGGACCAAGCTCTTCCC 477
```

361	QY	CGA	CA	CGGGTGTGAGCTTCTGTGGCGGCTTACGACGACTGCGCGGAGCGGGCGGACAGCGC	420
478	DB	CGA	CA	CGGGTGTGAGCTTCTGTGGCGGCTTACGACGGTGTGGCGGAGCGGGCGACAGCGC	537
421	QY	GGC	CTGCTGGACCTACTCCCTGCACTGTGGACATCACCGGATGGCATGAGAGCATCAAGGA	480	
538	DB	GGC	CTGCTGGACCTACTCCCTGCACTGTGGACATCACCGGATGGCATGAGAGCATCAAGGA	597	
481	QY	GGAGCTGAGGCGCCCTGGTCAAGGAGAAAGGGTGTGAACTCTCTTCTGGTCTTCATGGCATA	540		
598	DB	GGAGCTGAGGCGCCCTGGTCAAGGAGAAAGGGTGTGAACTCTCTTCTGGTCTTCATGGCATA	657		
541	QY	CAAGGA	CCGGTGCAGTGTGACGACGACGACAGATGTACGAGATCTTCAGCATCATCCGGGA	600	
658	DB	CAAGGA	CCGGTGCAGTGTGACGACGACGACAGATGTACGAGATCTTCAGCATCATCCGGGA	717	
601	QY	CCTGGGGCCCTTGGGCCCAAGGTGACGCTGAGAACCGGGACATCGTGGAGGAGGACGAA	660		
718	DB	CCTGGGGCCCTTGGGCCCAAGGTGACGCTGAGAACCGGGACATCGTGGAGGAGGACGAA	777		
661	QY	CGCGTGTCTGAGCTTCGGGCATCACTGGCGCCCGAGGGCCACGCTGCTCAGGCCACCCCGAGGA	720		
778	DB	CGCGTGTCTGAGCTTCGGGCATCACTGGCGCCCGAGGGCCACGCTGCTCAGGCCACCCCGAGGA	837		
721	QY	GGTGGAGGCTGAGGCGGTGTACCGAGCTGTCAACATCGCCAAAGCAGCAAACTGCCCGCT	780		
838	DB	GGTGGAGGCTGAGGCGGTGTACCGAGCTGTCAACATCGCCAAAGCAGCAAACTGCCCGCT	897		
781	QY	GTA	CGTCAACCAAGTGTATGAGCAAGGGGGCGGCCGACGCCATCGCTCAGGCCAAGCGCAG	840	
898	DB	GTA	CGTCAACCAAGTGTATGAGCAAGGGGGCGGCCGACGCCATCGCTCAGGCCAAGCGCAG	957	
841	QY	AGGGGTGTCTGTTTTGGGGAGGCCATCACCGCCAGCTGTGGGACCGACGTTTCACACTA	900		
958	DB	AGGGGTGTCTGTTTTGGGGAGGCCATCACCGCCAGCTGTGGGACCGACGTTTCACACTA	1017		
901	QY	CTGGAGCAAGAACTGGGGCAAGGCTTGACGCTTCGTCACTACATCACCCCTGTCAACCCAGA	960		
1018	DB	CTGGAGCAAGAACTGGGGCAAGGCTTGACGCTTCGTCACTACATCACCCCTGTCAACCCAGA	1077		
961	QY	CCCCACACCGGCAGACACCTCTCACTGTGTCTGTCCAGCGGGGACCTCCAGGTGACAGG	1020		
1078	DB	CCCCACACCGGCAGACACCTCTCACTGTGTCTGTCCAGCGGGGACCTCCAGGTGACAGG	1137		
1021	QY	CAGCGCCACTGCACCTTCCACCTGCCAGAGGGCTGTGGCAAGGACAACTTCGCGCT	1080		
1138	DB	CAGCGCCACTGCACCTTCCACCTGCCAGAGGGCTGTGGCAAGGACAACTTCGCGCT	1197		
1081	QY	GATCCCGAGGGCACCAACGGCATTTGAGAGGCGATGTTCGATGTCTGGGAGAAATGTGT	1140		
1198	DB	GATCCCGAGGGCACCAACGGCATTTGAGAGGCGATGTTCGATGTCTGGGAGAAATGTGT	1257		
1141	QY	GGCCTCTGGGGAAGATGAGCAGGAATGATTTGCTCGGGGTGACAGTACAAATGCTGCCAA	1200		
1258	DB	GGCCTCTGGGGAAGATGAGCAGGAATGATTTGCTCGGGGTGACAGTACAAATGCTGCCAA	1317		
1201	QY	AATCTTCAATTTTATCCCAAGGAAGGGCGAGTGGCTGTGGGCTCTCAGCGCTGACCTGGT	1260		
1318	DB	AATCTTCAATTTTATCCCAAGGAAGGGCGAGTGGCTGTGGGCTCTCAGCGCTGACCTGGT	1377		
1261	QY	CATATGGAACCCCAAGGCCACCAAGATCATCTCTGCCAAGACCCCAAACTCTGAACGTGGA	1320		
1378	DB	CATATGGAACCCCAAGGCCACCAAGATCATCTCTGCCAAGACCCCAAACTCTGAACGTGGA	1437		
1321	QY	GTACAAACATCTTCAGAGGAGTGGATGTCGCGGGGAGCGCTGCCCTGTGTGTCATTAAGTCAAGG	1380		
1438	DB	GTACAAACATCTTCAGAGGAGTGGATGTCGCGGGGAGCGCTGCCCTGTGTGTCATTAAGTCAAGG	1497		
1381	QY	CCGAGTGGCGCTTGAGGAGCGGAAGATGTTTGTCAACCCCGGGGCGGGCGCTTCGTC	1440		
1498	DB	CCGAGTGGCGCTTGAGGAGCGGAAGATGTTTGTCAACCCCGGGGCGGGCGCTTCGTC	1557		
1441	QY	TCGGAAAAACATTCGCCGACTTTGTCTACAGAGGATCAAAAGCTCGCAACAGGCTCGGCGGA	1500		

Db	1558	TCGGAATCATCCCGACTTTCTACAGAGATCAAGCTCGCAACAGGCTGGCGGA	1517
Qy	1501	GATCCACGGTGTCCCGGTGGGTGTATGACGGGCCGTCCACAGGTGATGTGCTGTC	1560
Db	1618	GATCCACGGTGTCCCGGTGGGTGTATGACGGGCCGTCCACAGGTGATGTGCTGTC	1677
Qy	1561	CAAGCCAGGAGTGGGGCTCCGGCCCGCGGTCTTCGCCACGACAGATCTCCGTGCTCC	1620
Db	1678	CAAGCCAGGAGTGGGGCTCCGGCCCGCGGTCTTCGCCACGACAGATCTCCGTGCTCC	1737
Qy	1621	TGTGCGCAACTTACATCAGTCTGGGGTTACGCTTATCTGGGTCTCAGGCTGATGACCAT	1680
Db	1738	TGTGCGCAACTTACATCAGTCTGGGGTTACGCTTATCTGGGTCTCAGGCTGATGACCAT	1797
Qy	1681	CGCCCGACGC	1690
Db	1798	CGCCCGACGC	1807

RESIST 5

RESOL 3  
AAD57506  
ID AAD57506 standard: cDNA: 2352 BP.

AX  
AC AAD57506:

XX DT 20-NOV-2003 (first entry)

XX  
DE Human enzyme (ENZM) CDNA #18

Human; enzyme; ENZM; immune disorder; infection; myocardial infarction; gene therapy; anaemia; acquired immune deficiency syndrome; infection; reproductive disorder; cardiovascular; eye; cell proliferation; cancer; AIDS; allergy; asthma; Addison's disease; diabetes; goitre; impotence; infertility; atherosclerosis; metabolic disorder; gene: ss.

XX  
QS  
Homo sapiens.

XX	Key	Location/Qualifiers
PH		
FT	CDS	94 . 1632

PN WO2003052075-A2.

XX  
26-JUN-2003XX  
PF 12-DEC-2002: 2002W0-IIS040161XX  
DB 14-DEC-2001: 2001118-034035

PR 20-DEC-2001; 2001US-034296

PR 22-JAN-2002; 2002US-035110

PA (INCY-) INCYTE GENOMICS INC.

PI Chawla NK, Lee SY, Ring HZ

PI Baughn MR, Hafalia AJA, Jin P, Swarnakar A, Li JX, Marquis JF,  
PI Lee S, Gorvad AE, Sprague WW, Becha SD, Elliott VS,

DR WPI; 2003-533016/50.

XX

PT for treating a disease associated with decreased expression or overexpression of ENZM e.g. cancer.

PS Claim 5; Page 322-323; 264pp; English.

The invention relates to human enzyme (ENZM) polypeptides and their corresponding polynucleotides. ENZM sequences are useful for preparing a composition for diagnosing or treating a disease or condition associated

CC with decreased expression or overexpression of functional ENZM. The  
CC disorders include immune disorders (anemia, allergy or asthma),  
CC infectious disorders (viral, fungal, parasitic or protozoal infection),  
CC immune deficiencies (acquired immune deficiency syndrome; AIDS),  
CC metabolic disorder (Addison's disease, diabetes or goitre), reproductive  
CC disorders (infertility or impotence), cardiovascular disorders  
CC (atherosclerosis or myocardial infarction), eye disorders and cell  
CC proliferative disorders (cancer). ENZM DNA is useful in gene therapy. The  
CC present sequence is human ENZM cDNA  
XX  
SQ Sequence 2352 BP; 471 A; 768 C; 711 G; 402 T; 0 U; 0 Other;

Query Match 78.08; Score 1318.4; DB 9; Length 2352;  
Best Local Similarity 89.3%; Pred. No. 1.4e-260;  
Matches 1509; Conservative 0; Mismatches 1; Indels 180; Gaps 1;

QY 1 GCGCCCTTACCAGACACCCAGGACGAGATCTCTCCAGGGCAAGAAAGCATCCC 60  
DB 63 GCGCCCTTACCAGACACCCAGGACGAGATCTCTCCAGGGCAAGAAAGCATCCC 122  
QY 61 CCGGATCACGAGTGACCGCCTTCTGATCAGAGGTGGGAGGATCGTGAATGACGACCAAGTC 120  
DB 123 CCGGATCACGAGTGACCGCCTTCTGATCAGAGGTGGGAGGATCGTGAATGACGACCAAGTC 182  
QY 121 CTTTACGCTGATGTGACGCTGGAAGATGGCTTGATATAAAACAAATCGGAGAAACCTCAT 180  
DB 183 CTTTACGCTGATGTGACGCTGGAAGATGGCTTGATATAAAACAAATCGGAGAAACCTCAT 242  
QY 181 CGTCCCTGGGGGATCAAGACCAATTGACGCCACCGGCTCATGCTCTCTCGTGGCGGT 240  
DB 243 CGTCCCTGGGGGATCAAGACCAATTGACGCCACCGGCTCATGCTCTCTCGTGGCGGT 302  
QY 241 TGAAGCTCACAAAGGCTGCAGATGCTCTCTGGGATGACACCGGCTGACGACTTCTG 300  
DB 303 TGAAGCTCACAAAGGCTGCAGATGCTCTCTGGGATGACACCGGCTGACGACTTCTG 362  
QY 301 TCAGGGACCAAGGACGCTAGCAGAGGAAACCAATGATCTTGAACCAAGCTTCTCC 360  
DB 363 TCAGGGACCAAGGACGCTAGCAGAGGAAACCAATGATCTTGAACCAAGCTTCTCC 422  
QY 361 CGACAGGGTGTAGCTGTGGCGCTACGAGCAGTGGGGAGCGGGCGGACGCGC 420  
DB 423 CGACAGGGTGTAGCTGTGGCGCTACGAGCAGTGGGGAGCGGGCGGACGCGC 482  
QY 421 GGCTGTGTGCACTACTCTCTGCACTGAGCATCACCCGATGGCATGAGAGCATCAAGGA 480  
DB 483 GGCTGTGTGCACTACTCTCTGCACTGAGCATCACCCGATGGCATGAGAGCATCAAGGA 542  
QY 481 GGAAGTGGAGGCTGTGTCAAGAGAAAGGTGTGAATCTCTTCTGTCTTCAATGGGATA 540  
DB 543 GGAAGTGGAGGCTGTGTCAAGAGAAAGGTGTGAATCTCTTCTGTCTTCAATGGGATA 602  
QY 541 CAAAGACCGGTGCGAGTGCAGCGACGACGATGTACGATCTTCAAGCATCTCCGGGA 600  
DB 603 CAAAGACCGGTGCGAGTGCAGCGACGACGATGTACGATCTTCAAGCATCTCCGGGA 662  
QY 601 CTTGGGGCTTGGCCCGAGGTGACGCTGAGAACGGGGACATCGTGGAGGAGGAGAGAA 660  
DB 663 CTTGGGGCTTGGCCCGAGGTGACGCTGAGAACGGGGACATCGTGGAGGAGGAGAGAA 722  
QY 661 GCGGTTGCTGGAGCTCGGATCTATGCCCCGAGGGCCAGCTGTCTCAGCCACCCGAGGA 720  
DB 723 GCGGTTGCTGGAGCTCGGATCTATGCCCCGAGGGCCAGCTGTCTCAGCCACCCGAGGA 782  
QY 721 GGTGGAGGCTGAGCGGTGTACCGAGCTGTACCATGCTCCAGCAGGCAAACTGCCGCT 780  
DB 783 GGTGGAGGCTGAGCGGTGTACCGAGCTGTACCATGCTCCAGCAGGCAAACTGCCGCT 842  
QY 781 GTACGTCACCAAGGTGATGAGCAAGGGGGCGGCGACCTCGCTCAGGCCAAGGCGAG 840  
DB 843 GTACGTCACCAAGGTGATGAGCAAGGGGGCGGCGACCTCGCTCAGGCCAAGGCGAG 902  
QY 841 AGGGGTGGTGTGTGTTGGGGAGGCCATCACCGCGAGCTGGGACCGGATTCACACTA 900

DB 903 AGGGGTGGTGTGTTTGGGGAGCCCATCACCGCAGCTGGGACCGACGCTTTCACACTA 962  
QY 901 CTGGAGCAAGAACTGGGGCCAAAGGCTGCAGCGCTTTCGTGCATCAACCCCTCTGCAACCCAGA 960  
DB 963 CTGGAGCAAGAACTGGGGCCAAAGGCTGCAGCGCTTTCGTGCATCAACCCCTCTGCAACCCAGA 1022  
QY 961 CCCACCAACGCGAGACCACTCACTCTCTGCTGTGTCTCAGCGGGAGACCTCCAGGTGACAGG 1020  
DB 1023 CCCACCAACGCGAGACCACTCACTCTCTGCTGTGTCTCAGCGGGAGACCTCCAGGTGACAGG 1082  
QY 1021 CAGGCGCCACTGCACCTTACCACTGCGCCAGAGAGGCTGTGGGCAAGCAAACTTCGCGCT 1080  
DB 1083 CAGGCGCCACTGCACCTTACCACTGCGCCAGAGAGGCTGTGGGCAAGCAAACTTCGCGCT 1142  
QY 1081 GATCCCGGAGGGCACCACCGCATTCAGGAGCGCATGTTCGATGTCTTGGGAGAAATGTGT 1140  
DB 1143 GATCCCGGAGGGCACCACCGCATTCAGGAGCGCATGTTCGATGTCTTGGGAGAAATGTGT 1202  
QY 1141 GGCTCTTGGGAGATGGACGAGAAATGAGTTCTGCGGGTGACCAAGTACAAATGTGCTCAA 1200  
DB 1203 GGCTCTTGGGAGATGGACGAGAAATGAGTTCTGCGGGTGACCAAGTACAAATGTGCTCAA 1262  
QY 1201 AATCTTCAATTTTACCCAAAGGAGGGGCGAGTGGCTGTGGGCTCTGACGCTGACCTGGT 1260  
DB 1263 AATCTTCAATTTTACCCAAAGGAGGGGCGAGTGGCTGTGGGCTCTGACGCTGACCTGGT 1322  
QY 1261 CATATGGAACCCCAAGGCCACCAAGATCATCTCTGCGCAAGACCAACANTCT 1320  
DB 1323 CATATGGAACCCCAAGGCCACCAAGATCATCTCTGCGCAAGACCAACANTCT 1373  
QY 1321 GTACAAACTCTTCAGGGAGTGGAGTGGCGGGAGCGCTTGGCTGTCTATAAGTCAAGG 1380  
DB 1374 ----- 1373  
QY 1381 CCGAGTGGCGCTGAGGAGCGGGAAGATGTTTGTCACCCCGGGGCGGGCGCTTTCGCTCC 1440  
DB 1374 ----- 1373  
QY 1441 TCGGAAAAATCTCCCGGACTTGTCTCAAGAGGATCAAGCTCGCAACAGGCTGGCGGA 1500  
DB 1374 -----GCTGGCGGA 1382  
QY 1501 GATTCACGCTGTGCCCGCTGTATGACGCGGCCCTGCACGAGGTGATGTGCTGCTGC 1560  
DB 1383 GATTCACGCTGTGCCCGCTGTATGACGCGGCCCTGCACGAGGTGATGTGCTGCTGC 1442  
QY 1561 CAAAGCAGGAGTGGGCTCTCCGCGCGGCTCTGCGCCAGGCAAGATCTTCGCTGCTCC 1620  
DB 1443 CAAAGCAGGAGTGGGCTCTCCGCGCGGCTCTGCGCCAGGCAAGATCTTCGCTGCTCC 1502  
QY 1621 TGTGCGCAACCTACATCAGTTCGGGGTTTCAAGCTTATCTGAGGTCTCAGGCTGATGACCAT 1680  
DB 1503 TGTGCGCAACCTACATCAGTTCGGGGTTTCAAGCTTATCTGAGGTCTCAGGCTGATGACCAT 1562  
QY 1681 CCGCCGACGC 1690  
DB 1563 CCGCCGACGC 1572

## RESULT 6

AAV60817

ID AAV60817 standard; cDNA; 1920 BP.

XX AAV60817;

AC AAV60817;

XX 08-DEC-1998 (first entry)

DT Mouse ULIP-4 coding sequence.

XX Mouse; Unc-33-like phospho-protein; ULIP; rat; neoplasm; tumorigenesis;

XX neurodegenerative disorder; diagnosis; ss.

XX





QY 1572 GTGGCGCTCCGGCCCGCGCGTCTGTCGCCAGGCAAGATCTCGTGCCTCTCTGTGCGCAACC 1631  
DB 1642 GTGGCACACAGGCCGCTGTCATCTCTGTTTCAAGCAAGATCTCAGTGCACACCGCTGCGCAACC 1701  
QY 1632 TACATCAGTCGGGCTTCAGGCTATCTGCGTCTCAGGCTGATGACCAATCGCCGAGC 1689  
DB 1702 TGACCAAGTCGGGCTTCAGGCTATCTGCTCTCAGGCTGAGATCAATTTGCCAGAGC 1759

RESULT 7  
ADK70708  
ID ADK70708 standard; DNA; 2470 BP.  
XX  
AC ADK70708;  
XX  
DT 06-MAY-2004 (first entry)  
XX  
DE Collapsin response mediator protein (CRMP) 3B DNA SeqID.  
XX  
KW gene; de; signal control molecule; collapsin response mediator protein;  
KW CRMP; nerve growth cone; axial future edge; neural-network formation;  
KW cell polarity; neurogenesis.  
XX  
OS Unidentified.  
XX  
PN JP200400094-A.  
XX  
PD 08-JAN-2004.  
XX  
PP 31-MAY-2002; 2002JP-00160853.  
XX  
PR 31-MAY-2002; 2002JP-00160853.  
XX  
PA (KAGA-) KAGAKU GIJUTSU SHINKO JIGYODAN.  
XX  
DR WPI; 2004-085208/09.  
DR P-PSDB; ADK70709.  
XX  
XX Novel N-terminal variant polypeptide of signal control molecule collapsin  
PT response mediator protein of nerve growth cone, useful for neural-network  
PT formation.  
PT  
XX Disclosure; Fig 18; 94pp; Japanese.  
PS  
XX This invention relates to a novel variant of an intracellular signal  
CC control molecule identified as the collapsin response mediator protein  
CC (CRMP) that works within the steering mechanism of the nerve growth cone  
CC in an axial future edge part. Specifically, it refers to an N-terminal  
CC variant that exhibits a deletion, substitution or addition of an amino  
CC acid in CRMP, as well as the appropriate monoclonal antibody. The present  
CC invention describes the CRMP variant as useful for neural-network  
CC formation and for the establishment of cell polarity during growth.  
CC Furthermore, it can be efficiently used during the neurogenesis process.  
CC This polynucleotide sequence is DNA encoding a CRMP isoform of the  
CC invention.  
XX  
SQ Sequence 2470 BP; 533 A; 685 C; 757 G; 495 T; 0 U; 0 Other;

Query Match 53.1%; Score 1067.2; DB 12; Length 2470;  
Best Local Similarity 77.9%; Pred. No. 4.4e-209;  
Matches 1301; Conservative 0; Mismatches 363; Indels 6; Gaps 1;

QY 27 GCAGGATGTCCTTCCAGGCGCAAGAAAGCATCCCGGATCAGAGTGACCGCTTCTGA 86  
DB 9 GGAAGATGTCGAACCCCGGCGAAGAGGCGCGCCCGCGGCAAGAGGACGCTGCTGA 69  
QY 87 TCAGAGTGGGAGGATGCTGAATGACGACAGTCCTTTTACGCTGATGTGCAAGTGAAG 146  
DB 69 TCAAGGGGGGAAATCGTCAACGACGACCGACGTCGTTTATGCGGACATTTATGTGAAG 128  
QY 147 ATGGCTTGATAAAACAAATCGGAGAAACCTCATGCTCCCTGGGGGATCAAGACCATG 206  
DB 129 ATGGGTTGATAAAACAGATTGGAGAGAAACCTGGCTGTGTTCCAGCGGGGTGAGGACGGTGG 188

QY 207 ACGCCCAACGCGCTGATGTCCTTCTCTGTGTCGCTTACGCTCCACACAAAGGCTGCAGATGC 266  
DB 189 ATGCTCTACGGGCAAGCTGGTGGTCCGGCGGACATCGATGTGCACACGCGCTGCAGATGG 248  
QY 267 CTGTCTCTGGGCATGACACCGGCTGACGATCTTGTGTAAGGSCACAAAGGCAAGCGCTAGCAG 326  
DB 249 CTGTGATGGGGATGGGCTCTGCTGATGATGACTTTTACCAAGGCAAGAGGCGCGCTGGCGG 308  
QY 327 GAGGAACCAACATGATCTTTGGACCAAGCTCTTCCCGCACACGCGTGTGAGCTGTGTCGCGG 386  
DB 309 GAGGAACCAACATGATGATGGAACCAAGTGTGCTGCCGAGGCTGGGACGAGCTGTGTCGCTG 368  
QY 387 CTTACGAGCAGTGGCGGAGCGGGCAGACGCGCGCTCTGCTCGGACTACTCTCCCTGCACG 446  
DB 369 CTTACGAGCAGTGGCGGCGAGCGTGTGACAGCAGAGGCGCTGCTCGGACTACGCGCTGCACA 428  
QY 447 TGGACATCACCCGATGGCATGAGAGCATCAAGGAGGAGCTGGAGGCGCTGCTGCAGAGA 506  
DB 429 TCGACATCCCGCTGGCATGAGAGCTTGGGAGGAGCTGGAGGCGCTGCTGTAAGGACA 488  
QY 507 AGGCTGTGAATCTCTTCTGCTCTTCATGSCATACAAAGGACCGGTGCCAGTGCAGGACA 566  
DB 489 AGGCTGTGAATCTCTTCTGCTGTTTTATGCGCTTACAAAGGACAGCTGCTGAGTGCAGCG 548  
QY 567 GCCAGATGTACGAGATCTTTCAGCATCATCCGGGACCTGGGGGCTTTGGGCCAGGTGCACG 626  
DB 549 CCCAGATGTATGAATATTTTCGATATTTTCGAGACCTTGGGGGCTAGTCCCAAGTGCACG 608  
QY 627 CTGAGAAACGGGAGCATCTGTGGAGGAGGAGCAGAAAGCGGTGTGTGGAGCTCGGATCACTG 686  
DB 609 CTGAAACCGGGAGCATCATCGAGGAGGAGCAGAAAGAGGCTGTGTGACATTTGGGATCACTG 668  
QY 687 GCGCCGAGGGCGACGTGCTCAGCCACCCGAGGAGGTGGAGGCTGAGGCGGTGTACCGAG 746  
DB 669 GGCAGAGGGGCGACGTCTGAGCGCGCTGAGAGGTTGAAGCAGAGGCGGTGTACCGTG 728  
QY 747 CTGTCAACCATCGCAAGCAGGCAAACTGGCCGCTGTACGCTCACAAGGTGATGAGCAAGG 806  
DB 729 CAATTACATAGCAAAACAGCAACTGCGCCCTTGTACGTCACCAAAATATAGCAAAA 788  
QY 807 GGGCGGCGGACGCGCATCGCTCAGGCGCAAGCGCAGAGGGTGTGTGTTTGGGAGGCCA 866  
DB 789 GTGCAGCAGATGTGTAGCTCAAGCAAGAGAAAGGCACTGTGTGTACGGGAGGCCA 848  
QY 867 TCACGCCAGCCTGGGACCGGAGGTTACACTTCTGAGGAGCAAGACTGGGCCAAGGCTG 926  
DB 849 TCACTGCCAGCTGGGTGCTGACGGCTCGACCTTCTGAGAGCAAGAACTGGGCCAAGGCTG 908  
QY 927 CAGCCTTCTGTCACATCACCCCTGTCAAGCCAGAGACCCCAAGGAGGAGGCTCAGCTT 986  
DB 909 CAGCCTTCTGTCACCTCACCCCGCTCAGCCCGGACCCCAAGGAGGAGGCTTCTCT 968  
QY 987 GCTTGTCTGTCAGCGGGGACCTTCAGGTGACAGGAGCGGCCCACTGCACCTTCCACATG 1046  
DB 969 CCTGCTGCTCTGTGGGAGCCTGCAAGTGGCAGGAGTGTCTCACTGACGTTCACTACCG 1028  
QY 1047 CCCAGAGGCTGTGGCAAGGACACTTCCGCTGATCCCGAGGAGGAGGAGGAGGATG 1106  
DB 1029 CCCAGAGGCGCGTGGGGAAGGACACTTCCACCTTCACTCCCGAGGAGGAGGAGGATG 1088  
QY 1107 AGGAGCGCATGTGATGTTGCGGAAATGTGTGCGCTCTGGGAAGATGGAGAGATG 1166  
DB 1089 AGAGCGGATGGCATATCTGGAGAGAGTGTGTGTCTCAGGGAAGATGATGAGAGC 1148  
QY 1167 AGTTCTGCGGTGACAGTACAAATGTCTGCCAAATCTTCAATTTTACCCAAAGAGG 1226  
DB 1149 ATTCTGTCGGGTGACAGCAGTAAATGTCTGCCAAATCTTTAACTGCTACCCCGAGAGG 1208  
QY 1227 GGGGAGTGTGCTGGGCTCTGAGCGCTGAGCTGTGATGGAACCCCAAGGAGGAGGAG 1286  
DB 1209 GGGGAATCGCGGTGGGCTCTGATGACAGACTTGTGCTTGTGGAAACCCCGAGGAGGAGG 1268



QY 1030 CTGCACCTTACCACTGCCCCAGAGGCTGTGGGCAAGCAACTTCGGCTGTATCCCGGA 1089  
Db 1293 CTGCACGTTTCACTACCGCCAGAGAGCGTGTGGGAGGACAACTTCACTCCCTCATCCCGGA 1352  
QY 1090 GGGCACCAACGGCAATGAGAGCGCATGTTCATGGTCTGGGAGAAATGTGTGGCTCTGG 1149  
Db 1353 GGGGACGATGGCATCGAGGAGCGGATGGCCATTAATCTGGGAGAAATGTGTGGCTCTGG 1412  
QY 1150 GAAGATGACGAGAAATGAGTTCTGGCGGTGACCAAGTACAAATCTGCCAAATCTTCAA 1209  
Db 1413 GAAGATGGATGAGAAACGATTTCTGGTGGCGGTGACCAAGTACAAATCTTCAA 1472  
QY 1210 TTTTACCAAGGAGGGGCGAGTGTGTGGGCTCTGACGCTGACTGTGTATATGAA 1269  
Db 1473 CTGTACCCAGGAGGGGCGAATCGCGTGGGCTCTGATGACACTTGTCTTGTGGAA 1532  
QY 1370 CCCCAGGCCACCAAGATCATCTCTGCCAGACCCCAATCTGAACTGAGGTGAGTACACAT 1329  
Db 1533 CCCCAGGGCTACAAAGTCACTCTCGGCAAAACCAATTTGATGTGGAGTACAAAT 1592  
QY 1330 CTTCAGGGAGTGGAGTGCCTGGGAGCGCTCTCGGTGTCTAATAGTCAGGCGCGAGTGGC 1389  
Db 1593 ATTTGAAGGCACGAGTGTACGCGCTCTCTGTGTGTCTCATCATGTCAGGGCAAGTGT 1652  
QY 1390 GCTGAGGACGGGAGATGTTTGTACCGCGGGGGCGGCTTGTCTCTCGGAAAC 1449  
Db 1653 TCTTGAGGACGGGAACTGTTTGTACCGAGGGCTCGGGGCGCTTCTGTTCCAGAAAGAC 1712  
QY 1450 ATTCGGGACTTCTCTACAGAGATCAAGCTCGCAACAGGCTGGGCGGAGTCCACGG 1509  
Db 1713 GTTCCCGATTTTGTATTAAAGAAATAAGCGGAGGAAACAGGCTGGCGAGGTGACGG 1772  
QY 1510 TGTGCGCGTGGGCTGTATGACGGGCGCGTCCAGAGGTGATGTGCTGCTGCAAGCCAGG 1569  
Db 1773 TGTCCACGCGGGCTGTATGACGGCGCGTCCAGATGTGCTCTCAGACCAAGTGT 1832  
QY 1570 GAGTGGGCTCCGCGC-----CGCGGCTCTGCGGAGGCAAGATCTCGTGTCTCTGT 1623  
Db 1833 GCCCACCATCGCGGCTCCAGGATCGGCGCTGCGCGGCAAGTGTGCGCGCGTCCGCT 1892  
QY 1624 GCGCAACTATACATCAGTGGGGTTTCACTTATCTGGGTCTCAGGCTGATGACCACTCGC 1683  
Db 1893 GCGCAACTGACAGTGGGGTTTCACTTATCTGGGTCTCAGGCTGATGACCACTCGC 1952  
QY 1684 CCGACAC 1690  
Db 1953 CCGGCGC 1959

## RESULT 9

ADJ84229  
ID ADJ84229 standard; DNA; 2947 BP.  
XX AC ADJ84229;  
XX XX  
XX 06-MAY-2004 (first entry)  
XX XX  
XX Malayan black rat CRMP-2 DNA which is a target for antisense therapy.  
XX XX  
XX antimicrobial; antinflammatory; cytostatic; infection; inflammation;  
XX KW tumour formation; antisense therapy; Malayan black rat; CRMP-2;  
XX KW collapse response mediator protein 2; ds; target; gene.  
XX XX  
XX Rattus rattus.  
XX XX  
XX WO2004003134-A2.  
XX XX  
XX 08-JAN-2004.  
XX XX  
XX 12-JUN-2003; 2003WO-US018481.  
XX XX  
XX 26-JUN-2002; 2002US-0392020P.  
XX XX

(ISIS-) ISIS PHARM INC.

Monia BP, Freier SM, Manoharan M, Gaarde WA;

WPI; 2004-083026/08.

P-PSDB; ADJ84279.

Decreasing levels of a preselected cellular mRNA in a cell, useful for treating or preventing a disease associated with a preselected cellular mRNA, comprises binding to the mRNA an antisense compound targeted to a splice site on the mRNA.

Example 19; SEQ ID NO 48; 174pp; English.

The invention relates to a novel method for decreasing levels of a preselected cellular mRNA in a cell or tissue comprising binding to the preselected cellular mRNA an antisense compound which is specifically hybridisable with a splice site on the mRNA and which is not a substrate for RNase H when bound to the RNA. The method of the invention has an antimicrobial, antinflammatory and cytostatic applications and may be useful for decreasing levels of a preselected cellular mRNA in a cell or tissue and thus for treating or preventing a disease or condition associated with a preselected cellular mRNA or with a preselected target cellular protein, particularly infection, inflammation or tumour formation. The current sequence is that of the Malayan black rat CRMP-2 (collapse response mediator protein 2) DNA of the invention which is a target for antisense therapy.

SQ Sequence 2947 BP; 666 A; 765 C; 769 G; 747 T; 0 U; 0 Other;

Query Match 58.0%; Score 980.2; DB 12; Length 2947;

Best Local Similarity 74.1%; Pred. No. 3.1e-191;

Matches 1240; Conservative 0; Mismatches 433; Indels 0; Gaps 0;

QY 18 CCCCCAGGAGCAGATGCTCTCCAGGCAAGAAAGCATCCCCCGATCAGAGTGACC 77

Db 164 CGCCCCCGGAGAGATGTCTTATCAGGGGAAGAAAAATATTCACGCATCAGAGGCATC 223

QY 78 GCCTCTGATCAGAGTGGGAGGATCGTGAATGACGACCACTCTTTTACGCTGATGTC 137

Db 224 GTCTTCTGATCAAGTGGCAAGATTGTGATGATGACCACTCTTCTATGACAGCATAT 283

QY 138 ACOTGGAAGATGGCTTTGATATAAAACAAATCGAGAAAACCTCATGCTCCCTGGGGCATCA 197

Db 284 ACATGGAAGATGGTTGATCAAGCAATAGGAGAAAACCTGATTGTGCCAGGAGGGTGA 343

QY 198 AGACCAATTGACGCCACCGGCTGATGCTCTTCTGTTGGCTTGAGCTGCACACAGGC 257

Db 344 AGACCATCGAAGCCCACTCCAGAAATGGTGTATCTCCCTGGAGGAATTGACACTCGCT 403

QY 258 TGCAGATGCTGTCTCTGGGCGATGACACCGCTGACGACTTCTGTCAGGGCACCAGGCAG 317

Db 404 TCCAGATGCCAGACCGAGGAATGACATCAGTGTATGATGATCTTCTCCAGGAGAACAGGCAG 463

QY 318 CGCTAGCAGAGGAACCAACCATGATCTTGGACCAACGCTCTTCCCGGACACGGGTGTGAGCC 377

Db 464 CCCTGCCCGAGGAACCAACCATGATCATCGACCATGTTCTGAGCCCGGAGCAAGCC 523

QY 378 TGTGTCGGGCTACGAGCAGTGGCGGAGGCGGCGGACAGCGGCTGTGCTGACTACT 437

Db 524 TATTGGACGCTTTGATTCAGTGGAGGAGTGGGCGGACAGCAAGTCTCTGCTGACTATT 583

QY 438 CCCTGCACGTTGGACATCACCCGATGCGATGAGAGCATCAAGGAGGAGCTGGAGGCCCTGG 497

Db 584 CGCTGCACGTTGGACATCACGGAGTGGCAACAGGGCATCCAGGAGGAGATGGAAGCTCTGG 643

QY 498 TCAAGGAGAGGGTGTGAATCTCTTCTCTGTTCTTCAATGGCATACAAGGACCCGGTGCAGT 557

Db 644 TGAAGGACCAAGGGGTAAACTCTCTCTCTGTTACATGCTTTTCAAGATCGGTTCCAGC 703

QY 558 GCAGCCACAGCCAGATGTACGAGATCTTCAGCATCATTCGGGACCTGCGGGGCTTGGCCC 617

Db 704 TGACGGATTCACAGATCTATGAAGTACTGAGCGTGTATCCGGGATATTGGTCCATAGCTC 763



Db 249 TCCAGATGCCCTGACAGGGAATGACATCCGCTGATGACTTCTCCAGGGAACCAAGGCGG 308  
Qy 318 CGCTAGCAGGAGGAACCAACATGATCTTTGGACCACTGCTTCCCGGACACAGGCTGAGCC 377  
Db 309 CCCTGGCCGGGGGAACCAACATGATCATGACCATGTTGCTTCCAGCCCGGGAGAGCC 368  
Qy 378 TGTGGGGCCCTACGAGAGTGGCGGAGCGGGCGGACAGCGCGCTGCTGCGACTACT 437  
Db 369 TATTGGCTGCTTTGATCAGTGGAGGAGTGGGCTGACAGCAAGTCTCTGCTGACTATT 428  
Qy 438 CCCTGCACGTGGACATCACCCGATGGCATGAGAGCATCAGGAGGAGCTGGAGCCCTGG 497  
Db 429 CGCTGCACGTGGACATCACAGTGGCAACAGGGAATCCAGGAGGAGATGGAAGCTCTGG 488  
Qy 498 TCAGGAGAGAGGGTGTGAATCTCTCTGCTGCTTCTCATGGGATACAAAGGACCGGTGCCAGT 557  
Db 489 TGAAGGACCAACCGGGTAACTCTCTCTGCTGCTGATGCTTTCAGAGTGCATTCGAGC 548  
Qy 558 GCAGCGACAGCAGATGATCAGAGATCTTCAGCATCATCCGGGACCTGGGGCCCTTGGCCC 617  
Db 549 TGACGGATTCCTCAGATCTATGAGTGTGAGCGGTGATCCGGGATATCGGTGCCATGCTC 608  
Qy 618 AGGTGCACGTGAGAACCGGGACATCGTGGAGGAGGAGCAGAGCGGTTGCTGGAGCTCG 677  
Db 609 AAGTCCACCGCAGAGAAATGGTGACATCATGCTGAGGCGACACAGCAGAGGATCCTGGATCTGG 668  
Qy 678 GCATCACTGGCCCCGAGGGCCACGTGCTCAGCCACCCCGAGGAGGTGGAGGCTGAGCGG 737  
Db 669 GCATCAGCGGCCCGAGGGACACGTGTTGAGCGCGCCAGAGAGGTGCGAGGCTGAGCTG 728  
Qy 738 TGTACCGAGCTGTACCATCGCCAAAGCAGGAGAACTCCCGCTGTAGCTCACCAGGTGA 797  
Db 729 TGAACCGGTCCATCACTATTGTCACCAACAGAGCACTGCTCTGTATGTCAACAAAGTGA 788  
Qy 798 TGAGCAAGGGGGCGCGGCGCATGCTCAGGCCAAGCCAGGAGGAGGGGTGCTGTTGTTG 857  
Db 789 TGCCCAAGAGTGGCGCTGAAGTCACTGCTCAGGACCGGAAGAGGGAACCTGCTGCTGATG 848  
Qy 858 GGGAGCCCATCACCGCCAGCGCTGGGCACCGACGCTTCCACTACTGGAGCAAGAACTGGG 917  
Db 849 GTGAGCCCATCAGCGCCAGCGCTGGGAGCTGATGGCTCTGATTAAGAGCAGAACTGGG 908  
Qy 918 CCAAGGTGAGCGCTTGTGTCATCACTACCCCTGTCAACCCAGACCCCAACAGCGGAGCC 977  
Db 909 CCAAGGTGCGGGCTTGTGTCATCCCTCCACCTTGGAGCCCGGACCAACCACTCCAGACT 968  
Qy 978 ACCTCACTGCTGCTGCTCAGCGGGGACCTCAGGTGACAGGAGCGGCCCACTGCACCT 1037  
Db 969 TTCTCAACTGCTGCTGCTGCTGAGAGACCTCAGAGTCACTGGCAGTGGCCCACTGCACCT 1028  
Qy 1038 TCACCACTGCGCAGAGGCTGTGGGCAAGGACAACTTTCGCGCTGATCCCGAGGGGACCA 1097  
Db 1029 TCAACACTGCGCAGAGGCTGTGGGAGAGGACAACTTTCACCTTGAATCCCGAGGGGACCA 1088  
Qy 1098 ACGGCAATGAGGAGCGCATGTCGATGCTCTGGAGAAATGTGGGCTCTGGGAAGATGG 1157  
Db 1089 ACGGCACTGAGGAGCGCATGCTGTCATTTGGGATAAAGCTGTGGTCACTGGGAAGATGG 1148  
Qy 1158 ACCAGATGATGCTGCTGCGGTGACCGATCAAAATGCTGCCAAAATCTTCAATTTTACC 1217  
Db 1149 ATGAGAATCAGTTTGTGGCTGTGACGAGCAACCAACGAGCCAAAGTCTTCAACTTTTACC 1208  
Qy 1218 CAAGGAAGGGCGAGTGGCTGTGGGCTCTGACGCTGACCTGCTCATATGGAACCCCAAGG 1277  
Db 1209 CCGGGAAGGTGCGATCTCGTGGGATCTGATGCTGACTTGGTCACTTGGGACCTTGACA 1268  
Qy 1278 CCACCAAGATCACTCTGCGCAAGACCCCAATCTGAACGTGGAGTACAACATCTTCGAGG 1337  
Db 1269 GTGTGAAGACCACTCTCTGCCAAGACACACACAGTGTCTTGGTACACATCTTTGAAG 1328  
Qy 1338 GAGTGAAGTGC CGGGAGCGCTTGGCGTGTGATTAAGTCAAGGCCGAGTGGCGCTCGAGG 1397

Db 1329 GCATGAGAGTGTGCGGGCTCCCCACTGGTGTGTCATCAGCCAGGGCAAGATTGTCTCGAGG 1388  
Qy 1398 ACCGGGAAGATGTTGTTCACCCCGGGCGCGCTTCGTCCCTCGGAAAACATTTCCCGG 1457  
Db 1389 ACGGCACACTTCAATGTCTCACTGAAGGCTCAGAGGCTACATTTCCCGGAGCCCTTCCTG 1448  
Qy 1458 ACTTTGTCTTACAAGAGGATCAAAAGCTGCGCAACAGGCTGGCGGAGATCCACGGTGTGCCCC 1517  
Db 1449 ACTTTGTGTACAAACGATCAAAAGCAAGGAGCAGGCTGGCTGAGCTGAGAGGGGTCCCTC 1508  
Qy 1518 GTGGGCTGTATGAGCGGCCCGTCCACGAGTGTATGCTGCTGCGCAAGCCAGGAGTGGCG 1577  
Db 1509 GTGGCTGTATGAGCGGACCGGTATGCGAGGTGTCTGTGACGCCCAAGACGCTGACTCCAG 1568  
Qy 1578 CTCCGGCCCGCGGCTCTGCGCCAGGCAAGATCTCCGTGCTCTGTGCGCAACCTACATC 1637  
Db 1569 CCTCATCAGCTAAGACATCCCTTCCAGCAGCAGGACCACTTCTCGGAACCTGACCC 1628  
Qy 1638 AGTCGGGTTTACGCTTATCTGGGTCTCAGGCTGATGATCAACATCGCCCGGACGC 1690  
Db 1629 AGTCGGATTTCAGTGTGCTGCTGCTCAGATTGACGACAACTTTCCCGCGCGC 1681

## RESULT 11

ABK91188  
ID ABK91188 standard; cDNA; 1829 BP.

XX ABK91188;

DT 05-NOV-2002 (first entry)

XX Human cDNA encoding Ulip2/CRMP2 protein.

XX Human; ss; gene; Ulip2; CRMP2; collapsin response mediator protein;  
Unc-33-like protein; neurodegenerative disease; Alzheimer's disease;  
paraneoplastic neurodegenerative disease; PND; myelination;  
demyelination; remyelination; myelin disorder; multiple sclerosis;  
autoimmune neurodegenerative disorder; HTLV-1 associated myelopathy;  
human T lymphocyte virus 1.

XX Homo sapiens.

XX Key Location/Qualifiers  
CDS 72..1793  
FT /\*tag= a  
FT /product= "Ulip2/CRMP2"

XX US2002119944-A1.

XX 29-AUG-2002.

XX 09-NOV-2001; 2001US-00986632.

XX 09-NOV-2000; 2000US-0246751P.

XX (AGUE/) AGUERA M.

XX (BELI/) BELIN M.

XX (CHAR/) CHARRIER E.

XX (HONO/) HONORAT J.

XX (RICA/) RICARD D.

XX (ROGE/) ROGEMOND V.

XX Aguera M, Belin M, Charrier E, Honorat J, Ricard D, Rogemond V;

XX WPI; 2002-627172/67.

XX P-PSDB; ABG32228.

XX Prevention or treatment of myelin disorders, such as multiple sclerosis,

XX by administering an agent selected from a Ulip/CRMP protein, a nucleic

XX acid coding for the protein, or an antibody directed against protein.

XX Claim 14; Page 17-18; 44pp; English.

The invention relates to a new method for prevention or treatment of myelin disorders, comprises administering to a patient an effective amount of an agent selected from a Ulp1 (Unc-33-like protein)/CRMP (collapsin response mediator protein) protein, a nucleic acid coding for Ulp1/CRMP, an antisense sequence capable of specifically hybridizing with the nucleic acid, an antibody directed against Ulp1/CRMP, or an aptamer capable of binding Ulp1/CRMP, and a pharmacologically acceptable carrier. Also included are methods of diagnosing a myelin disorder in a subject, identifying agents useful for the prevention or treatment of myelin disorders, using the Ulp1/CRMP protein/nucleic acids, agents capable of modulating the function or expression of the proteins (increasing or decreasing), and a method for identifying an endogenous agent as a therapeutic target for the prevention or the treatment of myelin disorders. The agents are useful for preventing or treating a myelin disorder such as multiple sclerosis or HTLV-1 (human T lymphocyte virus 1) associated myelopathy and neurodegenerative diseases, Alzheimer's disease, paraneoplastic neurodegenerative diseases (PND), autoimmune neurodegenerative disorder. Ulp1/CRMP proteins are involved in the pathogenesis of myelination, demyelination and remyelination. Antibodies to a Ulp1/CRMP protein are useful for diagnosing a myelin disorder. The present sequence encodes Ulp12/CRMP2 protein

Sequence 1829 BP; 426 A; 509 C; 511 G; 383 T; 0 U; 0 Other;

Query Match 57.1%; Score 965.4; DB 6; Length 1829;  
Best Local Similarity 73.6%; Pred. No. 3.1e-188;  
Matches 1230; Conservative 0; Mismatches 441; Indels 0; Gaps 0;

20 CCAGAGCAGGATGTCCTTCCAGGCAAGAAAGCATCCCCGGATFCAGAGTACCGC 79  
60 CCAGAGCAGGATGTCCTTATCAGGCGAAGAAATATTCACGCATCAGAGCATCGT 119  
80 CTTCTGATCAGAGTGGGAGTCTGTAATGACGACGATCTTTTACGCTGATGTCAC 139  
120 CTTCTGATCAAGAGGATGTAATTTGTAATGATGACGATGCTTATGACAGATATC 179  
140 GTGCAAGATGGCTGTATATAAATCGGAGAAACCTCATCGCTCCGCGGCATCAAG 199  
180 ATGGAAGATGGTGTATCAAGCAATAGGAAATCTGATGTGCGAGGAGATGTAAG 239  
200 ACCATTGACGCCACCGCTGATGGTCTTCTGGTGGGCTTCAAGCTTCAACAAGGCTG 259  
240 ACCATGAGGCCACTCCCGATGGTGTATCCCGAGGAATTCAGCTTCAACATCGTTTC 299  
260 CAGATGCTGCTGGGCATGACACCGCTGACGATCTTCTGAGGGCACCAGGACGCG 319  
300 CAGATGCTGATCAGGGAATGACGCTGCTGATGATTTCTTCAAGGAACCAAGGCGCC 359  
320 CTAGCAGGAGGAACCATGATCTTGGACACGCTCTTCCCGACACGCGGTGTAGCCTG 379  
360 CTGGCTGGGGAAACCATATGATATGACATGACATGCTTCTGAGCCTGGGCAAGCCTG 419  
380 CTGGCGGCTTACGAGCAGTGGCGGAGCGGGGAGCAGCGCGCTTCTGCGACTACTCC 439  
420 CTCGCTGCTTTGACCACTGGAGGAATGGGCGCAGCAGCAAGTCTCTGTGACTACTCT 479  
440 CTCGATGAGCATATCACCAGTGGCATGAGAGCATCAAGGAGAGCTGGAGGCGCTGGTC 499  
480 CTGATGTGGACATCAGCAGTGGCATTAAGGGATCCAGGAGGAGATGGAAGCGCTTGTG 539  
500 AAGGAGAGGGGTGAACTCTTCTGCTGCTTCTCATGGCATACAAGGACCGGTGCGAGTGC 559  
540 AAGGATCAAGGGGTAAATCTTCTCTGCTGTAATGGCTTCAAGATCGCTTCCAGCTA 599  
560 AGCGACAGCCAGATGTACAGAGATCTTTCAGCATCATCCGGACCTGGGGGCTTTGGCCAG 619  
600 ACGGATTCAGATTTATGAAGTACTGAGTGTGATCCGGGATATGGCGCCATAGCCCAA 659  
620 GTGACGCTGAGAACGGGACATCTGGAGAGGAGGAGAGCGGTGCTGGAGCTCGGC 679  
660 GTCCAGCAGAAAATGGCGACATCTTGCAGAGGAGCAGAGAGGATTCCTGGATCTGGGC 719  
680 ATCACTGGCCCCGAGGGCCACGCTGCTCAGCCACCCCGAGGAGGTGGAGGCTGAGGGCGTG 739

RESULT 12  
ACF25352  
ID ACF25352 standard; DNA; 4459 BP.  
XX  
AC ACF25352;

720 ATACGGGCGCCGAGGAGCATGTGCTGAGCCGACCTGAGGAGGTGAGGCCGAAGCCGCTG 779  
740 TACCGAGCTGTACCATGCGCAAGCAGGCAAACTGCCGCTGTAGCTCACCAGGTGATG 799  
780 AATCGTGCATCACCATGCGCAACAGACCAAACTGCCGCTGTATATCACCAGGTGATG 839  
800 AGCAAGGGGCGCGCGAGCCATCGCTCAGGCGCAAGCGCAGAGGGGTGCTGCTGTTGGG 859  
840 AGCAAAAGCTCTGCTGAGGTATCGCCCAAGGCAACGGAAGAGGAACTGTGTGTATGGC 899  
860 GAGCCCATCACCGCCAGCGCTGGGCGACCGAGGCTTCACTACTGGAGCAAGAACTGGGCC 919  
900 GAGCCCATCACTGCCAGCTTGGGAAACGAGCGGCTCCCATTTACTGGAGCAAGAACTGGGCC 959  
920 AAGGTGCGAGCTTGTGTACATCACCCCTGTCAACCCAGACCCACCGAGGAGCAAC 979  
960 AAGGTGCTGCTGCTTGTGTACCTCCCGACCTTTGAGCCCTGTATCCAACTCCAGACTTT 1019  
980 CTCACCTGCTGCTGTCAGCGCGGACCTTCAGGTGACAGGAGCGGCCACCTGCACTTC 1039  
1020 CTCAACTCCTGCTGCTGTGGAGACCTTCAGGTCAAGGCGAGTGCCTTGCAGGTTT 1079  
1040 ACCACTGCCAGAGGCTGTGGGCAAGGACAACTTGGCGCTGATCCCGAGGCGCAAC 1099  
1080 AACACTGCCAGAGGCTGTAGGAAAGGACAACTTCACTGATTCGGAGGCGCACTAAT 1139  
1100 GGCATTTAGGAGCGCATGTGATGTCTGGGAGAAATGTGTGGCTCTCTGGGAGATGGAC 1159  
1140 GGCATCTGAGGAGCGGATGTGCTGCTGAGCAAGGCTGTGTCTCACTGGGAGATGGAT 1199  
1160 GAGATGAGTCTGCTGCGGTGACAGTCAAAATGCTGCCAAATCTTCAATTTTACCCCA 1219  
1200 GAGAACAGTGTGTGGCTGTGACCGACCAATGAGCCAAAGTCTTCACTTTACCC 1259  
1220 AGGAAGGCGGAGTGTGGGCTGTGACGCTGACCTGCTCATATGGAACCCCAAGGCG 1279  
1260 CGGAAGGCGCATGCTGTGGGATCCGATCCGACCTGCTCATCTGGGACCCCGACAGC 1319  
1280 ACCAAGATCATCTGCGCAAGACCCCAATCTGAAGCTGGAGTACAACTCTCGAGGGA 1339  
1320 GTTAAACCATCTCTGCCAAGACACACACAGCTCTCTCGAGTACAACTCTTTGAAGC 1379  
1340 GTGGAGTCCGGGAGCGCTGCGGTGTGTCATAAGTCAAGGCGGAGTGGCTGGAGGAC 1399  
1380 ATGGAGTCCGGGCTCCCACTGCTGCTGATCAGCCAGGGAAGATGTCTTGGAGGAC 1439  
1400 GGAAGATGTTGTCAACCCCGGGCGGCGCTTCTGCTCCCTCGGAAAAACATTTCCCGGAC 1459  
1440 GGCACCTGCTGATGACCGAGGCTCTGGACGCTACATTTCCCGGAAGCCCTTCCCTGAT 1499  
1460 TTTGTCTAAGAGGATCAAGCTCGCAACAGCTGGCGGAGATCCACGCTGTGCCCGGT 1519  
1500 TTTGTTTCAAGCGTATCAAGCAAGGAGAGGCTGGCTGAGCTGAGAGGGGTTCCTCGT 1559  
1520 GGGCTGTATGAGGGGCGCTTCCACAGGTGATGCTGCTGCGCAAGCCAGGAGTGGCGCT 1579  
1560 GGCCTGTATGAGGACCCGCTGTGTGTAAGTGTCTGTGACGCCCAAGACAGTCACTCCAGC 1619  
1580 CCGGCCCGGCGCTTCTGCCCGAGGAGATCTCCGTGCTCTGTGCGCAACCTTACATCAG 1639  
1620 TCTCCGCGCAAGAGCTCTCTGCCAAGCAGAGGCGCCCACTGTCCGGAACCTGCGACAG 1679  
1640 TGGGGTTCAGGCTATCTGGGTCTAGGCTGTGATGACCAATGCCCGCGAGC 1690  
1680 TCTGGATTCAAGTGTGTGCTGCTGCTGATGATGATGATGATGATGATGATGATGATGAT 1730





```
QY 1520 GGGCTGTATGACGGGCGCTCCACGAGTGATGGTGCCTGCCAAGCCAGGAGTGGCGCT 1579
Db 1763 GGCTGTATGACGAGCTGTGTGAGTGTCTGTGAGCGCCCAAGACAGTCACTCCAGCC 1822
QY 1580 CCGGCCCGCGCGTCTCTGCCAGGCAAGATCTCCGTGCTCTCTGTGGCGCAACCTACATCAG 1639
Db 1823 TCCTCGGCCAAGAGCTCTCTGCCAAGCAGCAGGCCCCACCTGTCCGGAACCTGACACAG 1882
QY 1640 TCGGGGTTCAAGCCATCTCTGGGTCTCAGGCTCAGATGACACATCCGCCGACGC 1690
Db 1883 TCTGGATCAGTTGTGTGTGTCTCAGATTGATGACAAATCTCCCGCGCGC 1933

RESULT 13
ADC66333
ID ADC66333 standard; DNA; 4459 BP.
XX
AC ADC66333;
XX
DT 18-DEC-2003 (first entry)
XX
DE Human collapsin response mediator protein 2 gene 3' UTR sequence.
XX
KW neuroprotective; nootropic; neuroleptic; gene therapy;
KW human collapsin response mediator protein 2; neurodegenerative disease;
KW Alzheimer's disease; Down syndrome; schizophrenia; H-ras; ds.
XX
OS Homo sapiens.
XX
XX WC2003040320-A2.
XX
PD 15-MAY-2003.
XX
XX 04-NOV-2002; 2002MO-US035323.
XX
XX 08-NOV-2001; 2001US-00006911.
XX
XX (IS18-) ISIS PHARM INC.
XX
XX Gaarde WA, Watt AT;
XX WPI; 2003-449447/42.
XX
XX New compound, having a sequence targeted to a nucleic acid encoding human
XX collapsin response mediator protein 2, useful for preparing a composition
XX for treating neurodegenerative disease, e.g., Alzheimer's disease.
XX
XX Example 15; SEQ ID NO 11; 102pp; English.
XX
XX The invention relates to a new compound having a sequence comprising 8-50
XX bp targeted to a nucleic acid encoding human collapsin response mediator
XX protein 2 which specifically hybridizes with the nucleic acid encoding
XX human collapsin response mediator protein 2 and inhibits its expression.
XX The compound is useful for preparing a composition for treating
XX neurodegenerative disease, e.g., Alzheimer's disease, Down syndrome or
XX schizophrenia. This sequence represents the human collapsin response
XX mediator protein 2 gene 3' untranslated (3'UTR) sequence against which
XX the antisense oligonucleotides may be targeted.
XX
SQ Sequence 4459 BP; 1074 A; 1086 C; 1047 G; 1252 T; 0 U; 0 Other;
Query Match 57.0%; Score 963.8; DB 10; Length 4459;
Best Local Similarity 73.5%; Pred. No. 7.6e-188;
Matches 1229; Conservative 0; Mismatches 442; Indels 0; Gaps 0;
QY 20 CCAGGAGCAGGATGTCCTTCCAGGGCAAGAAAGCATCCCGGATCAGAGTGACCGC 79
Db 263 CCAGGAGAGAGATGTCCTTATCAGGGGAAGAAAATATCCAGCATCAGACGATCGT 322
QY 80 CTTCTCATCAGAGTGGGAGATCGTGAATGACGACCAAGTCTCTTTTACCGTGTATGTCAC 139
Db 323 CTTCTGATCAAGGAGGTAATAATTGTTAATGATGACCAGTCGTTCTATGCAGACATATAC 382
```

```
QY 140 GTGGAGATGGCTTGATATAAAACAAATCCGAGAAAACCTCATCGTCCCTGGGGGCATCAAG 199
Db 383 ATGGAGATGGGTGTGATCAAGCAAAATAGAGAAAATCTGATTTGTCAGGAGAGATGAAG 442
QY 200 ACCATTGACGCCACCGGCTGATGGTCTCTTCCTGTGGGTGGTGTGAGTGTCCACAAAGGCTG 259
Db 443 ACCATCGAGGCCCACTCCCGATGTGTATCCCGGAGGAATTTGACGTCCACACTCGTTTC 502
QY 260 CAGATGCTGTCTCTGGGCATGACACCGGCTGACGACTTCTGTACGGGCACCAAGGACGG 319
Db 503 CAGATGCTGTGATCAGGGAATGACGTCTGTGATGATTTCTTCCAAAGGAACCAAGGCGGCC 562
QY 320 CTAGCAGGAGGAACCAACCATGATCTTGGACACAGCTCTTCCCGGACACGGGTGTGAGCCCTG 379
Db 563 CTGGCTGGGGGAACCAATATGATCATTTGACACAGTGTTCCTGAGCCTGGGACAAGCCCTG 622
QY 380 CTGGGGGCTTACGAGCAGTGGCGGAGCGGGGAGCAGCGCGGCTCTGTGCGACTACTCC 439
Db 623 CTGCTGCTCTTTGACCAAGTGGAGGGAATGGGCGGACAGCAAGTCTCTGTGACTACTCT 682
QY 440 CTGCACTGTGACATCAACCGATGGCATGAGACATCAAGGAGGAGCTGAGGCGCTGGTC 499
Db 683 CTGCATGTGACATCAGCGAGTGGCATAGGGGATCCAGGAGGAGATGGAAGCGCTTGTG 742
QY 500 AAGGAGAGGGGTGTAACCTCTCTCTGCTCTTCTGTCATGCGCATACAAGACCGGTGCCAGTGC 559
Db 743 AAGGATCACGGGGTAAATTCCTTCTCTGCTGATCATGGCTTTCAAAGATCGCTTCCAGCTA 802
QY 560 AGCGACAGCAGATGTACAGATCTTTCAGCATCATCCGGGACCTTGGGGGCTTGGGCCAG 619
Db 803 ACGGATTGCCAGATTTATGAGTACTGAGTGTGATCCGGGATATTGGCGCCATAGCCCCAA 862
QY 620 GTGCACTGTGAAACGGGGACATCTGTGAGGAGGAGCAGAACGGTGTCTGTGAGCTCGGC 679
Db 863 GTTCCACGCAAGAAATGGCGACATCATTTGACAGGAGCAGCAGAGGATCTCTGATCTGGGC 922
QY 680 ATCATCTGGCCCGGAGGCCACGTCGTCTCAGCCACCCGAGGAGGTGGAGGCTGAGGCGGTG 739
Db 923 ATCAGGGGCCCGAGGGAACATGTGTCTGAGCCGACCTGAGAGGTTCAGGCGCCGAAGCCGTG 982
QY 740 TACCGAGCTGTCAACATCCCAAGCAGGCAAACTTGCCTGCTGTACGTCAACCAAGGTGATG 799
Db 983 AATGTGTCATCAACATCCCAACCAAGCACTTCCCTGTATATATACCAAGGTGATG 1042
QY 800 AGCAAGGGGGCGGCGACGCCATCTGCTCAGGCCAAGCGCAGAGGGGTGTGTTTGGG 859
Db 1043 AGCAAAAGCTCTGCTGAGGTCTATCGCCAGGCAAGGAAAGGAACTGTGTGTATGGC 1102
QY 860 GAGCCCATCACCGCAGCTGGGCAACCGGCTTACACTTCTGAGGCAAGCACTGGGGC 919
Db 1103 GAGCCCATCACTGCCAGCTTGGGAAACGGACGGCTCCCATTTCTTGGAGCAAGAACTGGGCC 1162
QY 920 AAGGCTGACGCTTCTGTACATCACACCCCTGTCAACCCAGACCCCAACACGCGACAGCAC 979
Db 1163 AAGGCTGCTGCTTGTGTACCTTCCACCTTGTAGGCCCTGTATCCACCACTCCAGACTTT 1222
QY 980 CTCACCTGCTGTCTGTCCAGCGGGACCTTCCAGGTGACAGGAGCGGCCCACTGTCACTTC 1039
Db 1223 CTCACCTGCTGTCTGTGTGAGACCTTCCAGGTCAAGGAGTGTGCCCATTTGACAGTTT 1282
QY 1040 ACCATGCTCCAGAAAGGCTGTGGCAGAGCAACTTTCGCTGATCCCGGAGGCAACCAAC 1099
Db 1283 AACACTGCCCAGAGGCTGTAGGAAAGGACAACTTCAACCTGATTCGGAGGGGCACAAT 1342
QY 1100 GGCAATTGAGAGCGCATGTGATGGTCTGGGAGAAATGTGTGGCTCTCTGGGAAGATGGAC 1159
Db 1343 GGCACTGAGAGCGGATGTCCGTCTATCTGGGACAGGCTGTGGTCACTGGGAAGATGGAT 1402
QY 1160 GAGAAATGATTTGCTCGCGGTGACCAAGTACAAATGTCTGCCAAAATCTTCAATTTTACCCA 1219
Db 1403 GAGAACCAAGTTGTGGCTGTGACCAAGCAATGACCAAAAGTCTTCAACCTTTACCCC 1462
```

Qy 1220 AGGAAGGGGAGTGGCTGTGGGCTTGACGCTGACCTGTGTCATATGAAACCCCAAGGCC 1279  
Db |||||  
Qy 1463 CGGAAGGCGCGCATTTGCTGTGGATCGGATCGCGACCTGTGTCATCTGGGACCCGACAGC 1522  
Db |||||  
Qy 1280 ACCAAGATCATCTCTGCCAAGACCCACAACTGAACTGAGTACACATCTTCGAGGA 1339  
Db |||||  
Qy 1523 GTTAAACCATCTCTGCCAAGACACACAACTGCTCTCGAGTACAACTCTTTGAAGGC 1582  
Db |||||  
Qy 1340 GTGAGTCCGCGGAGCGCTCCGCTGTCTCATAGTCAAGGCGCGAGTGGCGCTGGAGGAC 1399  
Db |||||  
Qy 1583 ATGGAGTCCGCGGCTCCCACTGTGTGTCATCAGCCAGGGAGATTTGCTCTGGAGGAC 1642  
Db |||||  
Qy 1400 GGAAGATGTTTGTACCCCGGGGGGCGGCTTCTCTCGGAAAAATTCCTCCGAGC 1459  
Db |||||  
Qy 1643 GGCACCTGCTATGTCACCGAAGGCTCTGCGACGCTTACATTCCTCCGGAAGCGCTTCCCTGAT 1702  
Db |||||  
Qy 1460 TTTGTCTACAGAGGATCAAGCTCGCAACAGGCTGGCGGAGATCCACGCTGTGCCCGT 1519  
Db |||||  
Qy 1703 TTTGTTTACAGCGGTATCAAGGCAAGGAGGAGGCTGGCTGAGCTGAGAGGGGTTCTCGT 1762  
Db |||||  
Qy 1520 GGGCTGTATGACGGGCGGCTCCACAGGCTGATGCTGCTGCCAAGCCAGGGAGTGGCGCT 1579  
Db |||||  
Qy 1763 GGCTGTATGACGGACCTGTGTGTAAGTGTCTGTGAGCGCCCAAGACAGTCACTCCAGCC 1822  
Db |||||  
Qy 1580 CCGGCGCGCGCTCTCCCGACAGGCAAGATCTCCGTGCTCTCTGTGCGCAACCTTACATCAG 1639  
Db |||||  
Qy 1823 TCCTCGGCCAAGAGCTCTCTCTGCCAAGCAGCAGGCGCCACCTGTGCGGAACCTGCAACAG 1882  
Db |||||  
Qy 1640 TCGGGTTACGCTATCTGGGTCTCAGGCTGATGACCAATCGCCGAGC 1690  
Db |||||  
Qy 1883 TCTGGATTGATGTTGTCTGGTCTCAGATTGATGACAACTTCCCGCGC 1933  
Db |||||

## RESULT 14

ADR98995 standard; DNA; 4459 BP.  
XX  
AC ADR98995;  
XX  
DT 02-DEC-2004 (first entry)  
XX  
DE Dihydropyrimidinase-like 2, DPYSL2, coding sequence, SEQ ID 1.  
XX  
XX Cytostatic; breast cancer; cancer; human; gene; ds;  
KW Dihydropyrimidinase-like 2; DPYSL2.  
KW  
XX Homo sapiens.  
XX  
XX WO2004078035-A2.  
XX  
XX 16-SEP-2004.  
XX  
XX 27-FEB-2004; 2004WO-US007268.  
XX  
XX 28-FEB-2003; 2003US-0450655P.  
XX  
XX (FARB ) BAYER PHARM CORP.  
XX  
XX Eveleigh D, Bigwood D;  
XX WPI; 2004-653556/63.  
XX P-PSDB; ADR99122.  
XX REFSEQ; NM\_001386.1.  
XX  
PT Diagnosing breast cancer comprises comparing the level of expression of  
PT genes or gene products in a first biological sample taken from a patient  
PT with that in a normal patient sample.  
XX  
XX Claim 2; SEQ ID NO 1; 53pp; English.  
XX  
XX The present invention relates to a method (M1) for diagnosing breast  
XX cancer in a patient. The method comprises comparing the level of  
XX expression of one or more genes or gene products in a biological sample

CC from the patient with that in a normal patient sample, where a difference  
CC in the gene expression in the first sample compared to that in the second  
CC sample is a diagnostic of the disease. Also claimed are: method (M2) for  
CC distinguishing between normal and disease tissues; method (M3) for  
CC monitoring the response of a breast cancer patient to treatment with an  
CC anti-cancer agent; method (M4) for identifying a compound for treating  
CC breast cancer; and an array for distinguishing between normal and disease  
CC tissues comprising two or more probes corresponding to genes selected  
CC from ADR98995-ADR99121 and comprising two or more polypeptides selected  
CC from ADR99122-ADR99248. In M1 and M2 the genes are selected from ADR98995  
CC -ADR99121 and the gene products are polypeptides selected from ADR99122-  
CC ADR99248. M1 is useful for diagnosing breast cancer. M2 and the array are  
CC useful for distinguishing between normal and disease tissue. M3 is useful  
CC for monitoring the response of a breast cancer patient to treatment with  
CC an anti-cancer agent. M4 is useful for identifying a compound for  
CC treating breast cancer. Note: The sequence data for this patent did not  
CC form part of the printed specification, but was obtained in electronic  
CC format directly from WIPO at ftp.wipo.int/pub/published\_pct\_sequences.  
XX  
SQ Sequence 4459 BP; 1074 A; 1086 C; 1047 G; 1252 T; 0 U; 0 Other;

Query Match 57.0%; Score 963.8; DB 13; Length 4459;  
Best Local Similarity 73.5%; Pred. No. 7.6e-188;  
Matches 1229; Conservative 0; Mismatches 442; Indels 0; Gaps 0;  
Qy 20 CCAGGAGCAGGATGTCCTTCCAGGGCAAGAAAGCATCCCCGGATCAGAGTGACCGC 79  
Db |||||  
Qy 80 CTTCTGATCAGAGTGGGAGGATCGTGAATGACGACAGTCTTTTACGCTGATGTGCAC 139  
Db |||||  
Qy 323 CTTCTGATCAAGAGGAGTAAATTTGTAATGATGACGACGTCGTTCTATGCAGACATATAC 382  
Qy 140 GTGGAGATGGCTTGATATAAAACAAATCGGAGAAACCTCATCTCTCCCTGGGGCATCAAG 199  
Db |||||  
Qy 383 ATGGAAGATGGGTGATCAAGCAATAGGAGAAATCTGATTTGCCAGGAGGATGAAG 442  
Qy 200 ACCATTGACGCGCCACCGGCTGTGTCCTTCTGTGGGCTTGACGTCCTCCACACAAAGGCTG 259  
Db |||||  
Qy 443 ACCATCGAGGCGCCACCTCCCGGATGGTGATCCCGGAGGAATTCACGCTCCACACTCGTTTC 502  
Qy 260 CAGATGCTGTCTCGGGCATGACACCGGCTGACGACTTCTGTGAGGGCACCAGAGGACGG 319  
Db |||||  
Qy 503 CAGATGCTGTATCAGGGAATGACGCTGTGCTGATGATTTCTTCCAGGAACCAAGGGCGCC 562  
Qy 320 CTAGCAGGAGGACCAACCATGATCTTTGGACCACTCTTCCCGCACACGGGTGTGAGCGTG 379  
Db |||||  
Qy 563 CTGGCTGGGGGAACCACTATGATCATTTGACCATGTTGTTCTGAGCTTGGGACAAGCCTG 622  
Qy 380 CTGGCGGCTTACGAGCAGTGGCGGAGCGGCGGACAGCGCGGCTCTGCTCGACTACTCC 439  
Db |||||  
Qy 623 CTGCTGCTTTGACCACTGGAGGGAATGGGCGCAGCAGCAAGTCTCTGTGTACTCT 682  
Qy 440 CTGCACTGGACATCACCCGATGGCATGAGAGCATCAAGAGAGCTGTGAGGCGCTGTGTC 499  
Db |||||  
Qy 683 CTGCATGTGGACATCAGCGAGTGGCATAGGGGATCCAGGAGGATGGAAGGCTGTGTG 742  
Qy 500 AAGGAGAGGGTGTGAATCTCTCTCTGCTTTTTCATGGCATACAGGACCGTTCGCGAGTGC 559  
Db |||||  
Qy 743 AAGGATCAAGGGGTAAATTTCTTCTCTGCTGATGATGGCTTTCAAGATCGCTTCCAGCTA 802  
Qy 560 AGCGACAGCCAGATGTACGAGATCTTTCAGCATCATCCGGGACCTGGGGGCTTGGGCCAG 619  
Db |||||  
Qy 803 ACGGATTCAGATTTTATGAGTACTGATGATGATCGGGATATTGGCGCCATAGCCAA 862  
Qy 620 GTGCACTGTAGAACCGGGACATCTGTGGAGGAGGAGCAGAGCGTGTGTGAGCTGCGC 679  
Db |||||  
Qy 863 GTCCAGCAGAAAATGGCGACATCAATTGAGAGGAGCAGCAGAGGATCTGTGATCTGGGC 922  
Qy 680 ATCACTGGCCCCCGAGGCGCACGTCGCTCAGCCACCCCGAGGAGGCTGAGGCGGTG 739  
Db |||||  
Qy 923 ATCAGCGGCCCCGAGGACATGTGTGAGCCGACCTGAGGAGGTGAGGCGCGAAGCGGTG 982



CC	to be tested for anti-neoplastic activity, determining a change in
CC	expression of at least one gene (I) of a signature gene set, where (I)
CC	comprises a sequence (S) selected from 847 sequences (given in ABU61664
CC	to ABU70110), or is at least 95% identical to (S), where a change in
CC	expression is indicative of anti-neoplastic activity. (I) has cytostatic
CC	activity and can be used in gene therapy. MI can be used for screening an
CC	anti-neoplastic agent, and can be used for producing a product which is
CC	the data collected with respect to the anti-neoplastic agent as a result
CC	of MI, and the data is sufficient to convey the chemical structure and/or
CC	properties of the agent. MI can be used in the treatment of cancer such
CC	as colon, breast, stomach, lung, thyroid, oesophageal, ovarian, kidney,
CC	prostate, or pancreatic cancer, adenocarcinoma, carcinoma, clear cell
CC	cancer, infiltrating ductal cancer, infiltrating lobular cancer, squamous
CC	cell carcinoma, neuroendocrine carcinoma, papillary carcinoma and Wilms'
CC	tumour
XX	
SQ	Sequence 5421 BP; 1390 A; 1269 C; 1209 G; 1553 T; 0 U; 0 Other;
Query Match	57.0%; Score 963.8; DB 6; Length 5421;
Best Local Similarity	73.5%; Pred. No. 7.8e-188;
Matches 1229; Conservative	0; Mismatches 442; Indels 0; Gaps 0;
QY	20
DB	1325
QY	80
DB	1385
QY	140
DB	1445
QY	200
DB	1505
QY	260
DB	1565
QY	320
DB	1625
QY	380
DB	1685
QY	440
DB	1745
QY	500
DB	1805
QY	560
DB	1865
QY	620
DB	1925
QY	680
DB	1985
QY	740
DB	2045

Search completed: September 24, 2005, 00:18:31  
Job time : 3124 secs

**THIS PAGE BLANK (MPTD)**

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 23, 2005, 23:10:32 ; Search time 25095 Seconds.  
(without alignments)  
2563.404 Million cell updates/sec

Title: US-09-367-496C-7  
Perfect score: 1690  
Sequence: 1 gcgcgcctaccagagacc.....atgaccacatgcgccgacgc 1690

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

EST.\*  
1: gb\_est1.\*  
2: gb\_est2.\*  
3: gb\_hic.\*  
4: gb\_est3.\*  
5: gb\_est4.\*  
6: gb\_est5.\*  
7: gb\_est6.\*  
8: gb\_ges1.\*  
9: gb\_ges2.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1683.6	99.6	1733	3	BC029280 Homo sapi
2	937.8	55.5	4239	3	BC046333 Mus muscu
3	928.6	54.9	1016	5	BX353556 BX353556
4	879	52.0	938	5	BQ956921 AGENCOURT
5	872.6	51.6	1075	5	BM927995
6	829.4	49.1	1899	3	AK077000 Mus muscu
7	823.2	48.7	857	5	BX441221 BX441221
8	810.4	48.0	3161	3	AK090197 Mus muscu
9	804.6	47.6	963	5	BX374512 BX374512
10	795.8	47.1	1105	5	BM928281 AGENCOURT
11	795	47.0	1058	5	BQ068024 AGENCOURT
12	793.2	46.9	1050	5	BM920859 AGENCOURT
13	790	46.7	1042	5	BX442264 BX442264
14	775.4	45.9	900	5	BX444582 BX444582
15	772.4	45.7	1680	9	AY419150 Mus muscu
16	763.8	45.2	1680	9	AY419148 Homo sapi
17	763.6	45.2	1100	4	BM547750 AGENCOURT
18	745.4	44.1	1120	4	BM805653 AGENCOURT
19	744.6	44.1	828	5	BX418351 BX418351
20	743.2	44.0	1120	5	BM925929 AGENCOURT
21	738.8	43.7	1069	5	BM927042 AGENCOURT
22	719.4	42.6	966	5	BM925798 AGENCOURT
23	719.4	42.6	1680	9	AY419149 Pan trogl
24	716.4	42.4	887	4	BI524483 603051613

c

## ALIGNMENTS

RESULT 1  
LOCUS BC029280 1733 bp mRNA linear HTC 01-MAY-2002  
DEFINITION Homo sapiens, clone IMAGE:4931416, mRNA.  
ACCESSION BC029280  
VERSION BC029280.1 GI:20379613  
KEYWORDS HTC.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 1733)  
AUTHORS Strausberg, R.  
TITLE Direct Submission  
JOURNAL Submitted (01-MAY-2002) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA

REMARK NIH-MGC Project URL: http://mgc.nci.nih.gov  
COMMENT Contact: MGC help desk  
Email: cgabs-rc@mail.nih.gov  
Tissue Procurement: David N. Louis, M.D.  
CDNA Library Preparation: Life Technologies, Inc.  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Baylor College of Medicine Human Genome Sequencing Center

Center code: BCM-HGSC  
Web site: http://www.hgsc.bcm.tmc.edu/cdna/  
Contact: angbcm.tmc.edu  
Gunnaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Hale, S.M., Yoon, V.S., Kowis, C.R., Lawrence, S., Martin, R.G., Muzny, D.M., Richards, S., Gibbs, R.A.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov  
Series: IRAK Plate: 29 Row: n Column: 24  
This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 11321616  
This clone has the following problem: no polyA-tail.

Location/Qualifiers  
1. 1733  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:4931416"  
/tissue\_type="Brain, anaplastic oligodendroglioma with 1p/19q loss"

FEATURES

source

1000000000

1000000000

1000000000

1000000000

1000000000

1000000000

1000000000

1000000000

1000000000

1000000000

1000000000

1000000000

1000000000

1000000000

1000000000

1000000000

1000000000

1000000000

1000000000

1000000000

1000000000

1000000000

1000000000

1000000000

1000000000

1000000000

1000000000



/clone_lib="NCI_CGAP_Brn67"	
/lab_host="DHI08"	
/note="Vector: pCMV-SPORT6"	
Query Match 99.6%; Score 1683.6; DB 3; Length 1733;	
Best Local Similarity 99.8%; Pred. No. 0;	
Matches 1686; Conservative 0; Mismatches 4; Indels 0; Gaps 0;	
QY	1 GCCTCCCTTACAGAGACCCCGAGGACAGAGATGCTTCCAGGGGCAAGAAAAGCATCC 60
Db	8 GCCTCCCTTACAGAGACCCCGAGGACAGAGATGCTTCCAGGGGCAAGAAAAGCATCC 67
QY	61 CCGGATCAGAGTGACCGCTTCTGATCAGAGTGCGGAGGATCGTGAATGACGACCATC 120
Db	68 CCGGATCAGAGTGACCGCTTCTGATCAGAGTGCGGAGGATCGTGAATGACGACCATC 127
QY	121 CTTTACGCTGATGTCACGTGGAAGATGCTTGAATAAAACAAATCGGAGAAACCTCAT 180
Db	128 CTTTACGCTGATGTCACGTGGAAGATGCTTGAATAAAACAAATCGGAGAAACCTCAT 187
QY	181 CGTCCCTGCGGGCATCAAGACCATTTGACGCCACCGCTGATGGTCTTCTGCTGGCGT 240
Db	188 CGTCCCTGCGGGCATCAAGACCATTTGACGCCACCGCTGATGGTCTTCTGCTGGCGT 247
QY	241 TGAGCTCCACACAGGCTGAGATGCTGCTGCGGCATGACACCGGCTGACGACTTCTG 300
Db	248 TGAGCTCCACACAGGCTGAGATGCTGCTGCGGCATGACACCGGCTGACGACTTCTG 307
QY	301 TCAGGGACCAAGCGAGCGCTAGCAGGAGGAACACCATGATCTTGGACCAACGCTTCCC 360
Db	308 TCAGGGACCAAGCGAGCGCTAGCAGGAGGAACACCATGATCTTGGACCAACGCTTCCC 367
QY	361 CGACACGGGTGTAGCTGCTGCGGCTTACGACGAGTGTGCGGAGCGGCGGACAGCGC 420
Db	368 CGACACGGGTGTAGCTGCTGCGGCTTACGACGAGTGTGCGGAGCGGCGGACAGCGC 427
QY	421 GGCCTGTGCGACTACTCTCCGTCAGCTGGACATCACCCGATGCGATGAGCATCAAGGA 480
Db	428 GGCCTGTGCGACTACTCTCCGTCAGCTGGACATCACCCGATGCGATGAGCATCAAGGA 487
QY	481 GGAAGTGGAGCCCTGTCTCAAGAGAGAGGTGTGAACCTCTCTGCTTCTCATGGCAT 540
Db	488 GGAAGTGGAGCCCTGTCTCAAGAGAGAGGTGTGAACCTCTCTGCTTCTCATGGCAT 547
QY	541 CAAGGACCGGTGTCAGTGACGACAGCCAGATGTACGAGATCTTACGACATCATCCGGGA 600
Db	548 CAAGGACCGGTGTCAGTGACGACAGCCAGATGTACGAGATCTTACGACATCATCCGGGA 607
QY	601 CCTGGGGGCTTGGCCGAGTGCACGCTGAGACCGGGACATCGTGGAGGAGGAGCAGAA 660
Db	608 CCTGGGGGCTTGGCCGAGTGCACGCTGAGACCGGGACATCGTGGAGGAGGAGCAGAA 667
QY	661 GCGGTGTGCTGAGCTCGGACATCACTGCGCCCGAGGGCCAGCTGCTCAGCCACCCCGAGGA 720
Db	668 GCGGTGTGCTGAGCTCGGACATCACTGCGCCCGAGGGCCAGCTGCTCAGCCACCCCGAGGA 727
QY	721 GGTGGAGGCTGAGCGCGTGTACCGAGCTGTACCATCGCCAGCAGGCAAACTGCCCGCT 780
Db	728 GGTGGAGGCTGAGCGCGTGTACCGAGCTGTACCATCGCCAGCAGGCAAACTGCCCGCT 787
QY	781 GTACGTCACCAAGGTGATGAGCAAGGGGGCGGCGAGCCATCGCTCAGGCGCAAGCGCAG 840
Db	788 GTACGTCACCAAGGTGATGAGCAAGGGGGCGGCGAGCCATCGCTCAGGCGCAAGCGCAG 847
QY	841 AGGGGTGGTGTGTTTGGGAGGCCATCACCGCAGCGCTGGGACCGACCGGTTCACTA 900
Db	848 AGGGGTGGTGTGTTTGGGAGGCCATCACCGCAGCGCTGGGACCGACCGGTTCACTA 907
QY	901 CTGGAGCAAGAACTGGGCCAAGGCTGACGCTTGTGTCATCATACCCCTGTCAACCCAGA 960
Db	908 CTGGAGCAAGAACTGGGCCAAGGCTGACGCTTGTGTCATCATACCCCTGTCAACCCAGA 967
ORIGIN	
QY	961 CCCACCCAGCGCAGACCCACTCCTCCTGCTTGTCTGTCAGCGGGGAGCTCCAGGTGACAG 1020
Db	968 CCCACCCAGCGCGGACCCACTCCTCCTGCTTGTCTGTCAGCGGGGAGCTCCAGGTGACAG 1027
QY	1021 CAGCGCCCACTGCACTTCCACCACTGCCCAGAGGCTGTGGGCAAGGACAACTTCGCGCT 1080
Db	1028 CAGCGCCCACTGCACTTCCACCACTGCCCAGAGGCTGTGGGCAAGGACAACTTCGCGCT 1087
QY	1081 GATCCCGGAGGACCAACCGCATTTGAGGAGCGCATGTGATGCTTGGGAGAAATGTGT 1140
Db	1088 GATCCCGGAGGACCAACCGCATTTGAGGAGCGCATGTGATGCTTGGGAGAAATGTGT 1147
QY	1141 GGCCTCTGGGAGATGAGACGAGAAATGAGTTCTGCGGCTGACAGTACAAATGCTGCCAA 1200
Db	1148 GGCCTCTGGGAGATGAGACGAGAAATGAGTTCTGCGGCTGACAGTACAAATGCTGCCAA 1207
QY	1201 AATCTTCAATTTTATCCCAAGGAGGCGAGTGTGGCTCTGACGCTGACCTGGT 1260
Db	1208 AATCTTCAATTTTATCCCAAGGAGGCGAGTGTGGCTCTGACGCTGACCTGGT 1267
QY	1261 CATATGGAACCCCAAGGACCAACGATCATCTCTGCAAGACCAACATCTGAACTGGA 1320
Db	1268 CATATGGAACCCCAAGGACCAACGATCATCTCTGCAAGACCAACATCTGAACTGGA 1327
QY	1321 GTACACATCTTTCGAGGGAGTGGAGTGTGCGGGAGGCTGCGCTGATTAAGTCAGG 1380
Db	1328 GTACACATCTTTCGAGGGAGTGGAGTGTGCGGGAGGCTGCGCTGATTAAGTCAGG 1387
QY	1381 CCGAGTGGCGCTTGGAGGACCGGGAAGATGTTTGTCAACCCCGGGGGCGGCTTCTGCTCC 1440
Db	1388 CCGAGTGGCGCTTGGAGGACCGGGAAGATGTTTGTCAACCCCGGGGGCGGCTTCTGCTCC 1447
QY	1441 TCGGAAAAATTCCTCGGACTTTTGTCTCAAGAGGATCAAAAGCTCGCAACAGGCTGGCGGA 1500
Db	1448 TCGGAAAAATTCCTCGGACTTTTGTCTCAAGAGGATCAAAAGCTCGCAACAGGCTGGCGGA 1507
QY	1501 GATCCACGGTGTGCGCGGCTGTATGACGCGGCGGCTCCACGAGGTGATGCTGCTGC 1560
Db	1508 GATCCACGGTGTGCGCGGCTGTATGACGCGGCGGCTCCACGAGGTGATGCTGCTGC 1567
QY	1561 CAAAGCAGGAGTGGCGCTCCGCGCGCGCTCCGCGCGCTCCGCGCGGCAAGATCTCCGTGCTCC 1620
Db	1568 CAAAGCAGGAGTGGCGCTCCGCGCGCGCTCCGCGCGCTCCGCGCGGCAAGATCTCCGTGCTCC 1627
QY	1621 TGTGCGCAACTACATCAGTCCGGGTTGAGCTATCTGGGCTCAGGCTGATGACACAT 1680
Db	1628 TGTGCGCAACTACATCAGTCCGGGTTGAGCTATCTGGGCTCAGGCTGATGACACAT 1687
QY	1681 CGCCCGACGC 1690
Db	1688 CGCCCGACGC 1697
RESULT 2	
BC046333	
LOCUS	
DEFINITION	
Mus musculus dihydropyrimidinase-like 4, mRNA (cdna clone	
IMAGE:6492897), with apparent retained intron.	
ACCESSION	
BC046333	
VERSION	
BC046333.1 GI:28374377	
KEYWORDS	
SOURCE	
Mus musculus (house mouse)	
ORGANISM	
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;	
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.	
REFERENCE	
1 (bases 1 to 4239)	
Auerbach, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,	
Klauser, R.D., Collins, F.S., Wagner, L., Shennan, C.M., Schuler, G.D.,	
Hopkins, R.F., Jordan, H., Moore, K.H., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,	
Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,	
Stapleton, M., Soares, M.B., Bonaldo, M.P., Casavant, T.L.,	
Scheetz, T.E., Brownstein, M.J., Udwin, T.B., Toshiyuki, S.,	

Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J., Abramson, R.D., Mullahy, S.J., Bosak, S.A., McEwan, P.J., McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S., Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W., Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Fahney, J., Helton, E., Kettman, M., Madan, A., Rodriguez, S., Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y., Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butterfield, Y.S., Krzywinski, M.I., Skalek, U., Smailus, D.E., Schnerch, A., Schein, J.E., Jones, S.J., and Marra, M.A.

Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences

Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

2388257

12477932

2 (bases 1 to 4239)

Strausberg, R.

Direct Submission

Submitted (31-JAN-2003) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA

NIH-MGC Project URL: <http://mgc.nci.nih.gov>

Contact: MGC help desk

Email: [cgaps-remail.nih.gov](mailto:cgaps-remail.nih.gov)

Tissue Procurement: The Cepko Laboratory

cDNA Library Preparation: Life Technologies, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILNL)

DNA Sequencing by: National Institutes of Health Intramural Sequencing Center (NISC), Gaithersburg, Maryland

Web site: <http://www.nisc.nih.gov/>

Contact: [nisc.mc@nhgri.nih.gov](mailto:nisc.mc@nhgri.nih.gov)

Akhter, N., Ayèle, K., Beckstrom-Sternberg, S.M., Benjamin, B., Blakesley, R.W., Bouffard, G.G., Breen, K., Brinkley, C., Brooks, S., Dietrich, N.L., Granite, S., Guan, X., Gupta, J., Haghighi, P., Hansen, N., Ho, S.-L., Karlins, E., Kwong, P., Laric, P., Legaspi, R., Maduro, Q.L., Masiello, C., Maskeri, B., Mastrian, S.D., McCloskey, J.C., McDowell, J., Pearson, R., Stantripop, S., Thomas, P.J., Touchman, J.W., Turgeon, C., Vogt, J.L., Walker, M.A., Wetherby, K.D., Wiggins, L., Young, A., Zhang, L.-H., and Green, E.D.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/ILNL at: <http://image.llnl.gov>

Series: IRAC Plate: 100 Row: o Column: 8

This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 6753677

This clone has the following problem: retained intron.

Location/Qualifiers

1. 4239

/organism="Mus musculus"

/mol\_type="mRNA"

/db\_xref="taxon:10090"

/clones="IMAGE:6492897"

/tissue\_type="Eye, retina, mouse strain C57Bl/6"

/clone\_lib="NIH\_MGC\_94"

/lab\_host="DH10B"

/note="Vector: pCMV-SPORT6"

FEATURES

source

ORIGIN

Query Match 55.5%; Score 937.8; DB 3; Length 4239;

Best Local Similarity 87.1%; Pred. No. 9.1e-200;

Matches 1029; Conservative 0; Mismatches 152; Indels 0; Gaps 0;

Qy 509 GGTGTGAACCTCTTCCTGGTCTTCATGGCATACAAAGACCGGTGCCAGTGACGACAGC 568

Db 2153 GGTGTGAACCTCTTCCTGGTCTTCATGGCATACAAAGACCGGTGCCAGTGACTGACGGT 2212

Qy 569 CAGATGTACGAGATCTTCAGCATCATCGGACCTCGGGGCTTGGCCCGAGGTGCAGCT 628

Db 2213 CAGATATATGAATCTTCAGCCTCATCGGACCTCGGGAGCTGTGGCCCGAGGTGCACGCA 2272

Qy 629 GAGAACGGGGACATCGTGGAGGAGGAGCAGAACGGTTCCTGGAGCTCGGCATCCTGCGC 688

Db 2273 GAAATGGGGACATCGTGGAGGAGGAGAACAGAGCGCTGCTGGAGCAAGGCATCCTGCT 2332

Qy 689 CCCGAGGGCCACGTGCTCAGCCACCCCGAGGAGGTGAGGGCTGAGGGGTGTACCGAGCT 748

Db 2333 CCTGAGGGGCCATGCTGCTCAGCCACCCAGAGAGGTAGAGCCCGAGGTGTGTACAGAGCA 2392

Qy 749 GTCAACATCGCCAAAGCAGGCGAACTGCCCGCTCTAGCTCACCAGGTGATGAGCAAGGG 808

Db 2393 GTCAACATTCGCAAGCAGGCGCACTGCCCACTATAGCTCACCAGGTGATGAGCAAGGG 2452

Qy 809 GCGGCCGACGCTCATCGCTCAGGCCAAGCGCAGAGGGTGTCTGTTTGGGGAGCCCATC 868

Db 2453 GCAGCTGACATGTTGCTCCCAAGCGCAGGGGGTGTGCTCTTTTGGGGAACCTATTC 2512

Qy 869 ACCGCGAGCTGGGCAACCGACGTTTCACTACTCTGGAGCAAGAACTGGGCGCAAGGCTGCA 928

Db 2513 ACTGCCAGGCTGGGCACTGATGGCTCACACTACTCTGGAGCAAGAACTGGGCGCAAGGCTGCA 2572

Qy 929 GCCTTGTGTCATCACCCTCTGTCAACCCAGACCCCAACCGGACCTTACTCTGACGACCACTCCTCT 988

Db 2573 GCCTTGTGTCATCACCCTCTGTCAACCCGACCTTACTCTGACGACCACTCCTCTCT 2632

Qy 989 TTGCTGTCCAGCGGGGACCTTCCAGGTGACAGGCGAGCCCACTGCACTTCACTGCTGC 1048

Db 2633 CTGCTGTCCAGTGGGACCTTCCAGGTGACAGGCGAGTGCCTGCTGCTGCTGCTGCTGCT 2692

Qy 1049 CAGAGGCTGTGGGCAAGGACAACTTCCGCTGATCCCGAGGGCACCACGCGCATTTGAG 1108

Db 2693 CAGAGGCTGTGGGCAAGGACAACTTCCGCTGATCCCGAGGGAGTCAACCGTATAGAA 2752

Qy 1109 GAGCGCATGTCCAGTGTCTGGGAGAAATGTGGCTCTTGGGAGAGTGAACGAGATGAG 1168

Db 2753 GAGCGCATGTCTGTGTCTGGGAGAAATGTGGCTCTGAGGAGAAATGAGCAGAAATGAG 2812

Qy 1169 TTGCTCGCGGTGACCAAGTCAAAATGTGCCAAATCTTCAATTTTACCCAGGAGAGGG 1228

Db 2813 TTGTTGGCGTGAACAGCAAAATGTGCCAAATCTTCAATTTTACCCAGGAGAGGG 2872

Qy 1229 CGAGTGGCTGTGGGCTCTGACGCTGACCTGCTGATATGNAACCCCAAGGCCACCAAGATC 1288

Db 2873 CGTGTGGCGTGGGCTCTGATGCTGACCTGCTGATCTGGAACCCCAAGGCCACCAAGATC 2932

Qy 1289 ATCTGTCCAAAGACCCACCAATCTGAACGTGAGTACAAATCTTCCAGGAGTGGAGTGC 1348

Db 2933 ATCTGTCCAAAGACCCATACCTGAATGTAGAGTACAAATCTTCCAGGAGTGGAGTGC 2992

Qy 1349 CGGGAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1408

Db 2993 CGAGGAGTGGCCACCGTGGTCTAAGTCAAGGCGAGTGGTGGTGGTGGTGGTGGTGGTGGT 3052

Qy 1409 TTTGTCACTCCCGGG 1468

Db 3053 CTTGTCACTCCAGGGGGTGGCGGCTTCAATTCCTCCCGGAGACGTTCCCGGAGCTTTGTCTAT 3112

Qy 1469 AAGAGGATCAAAAGCTGCAACAGGCTGGCGGAGATCCAGGCTGCTGCCCGCTGGCTGTAT 1528

Db 3113 AAGAGGATCAAAAGCTGCAACAGGCTGAGGAGATCCAGGCTGCTGCCCGCTGGCTGTAT 3172

Qy 1529 GACGGGCGGCTCCAGAGGTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1588

Db 3173 GACGGGCGGCTGTCATGATGATGTTTACCTGCCAAGCCAGGAGTGGCAACACAGGCGCGT 3232

Qy 1589 GCGTCTGCCAGGCAAGATCTCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1648

Db 3233 GCATCTGTTTCCAGGCAAGATCTCAGTGGCCACCGCTGCGCAACCTGACCAAGTGGGGTTC 3292

Qy 1649 AGCCTATCTGGGTCTCAGGCTGATGACCATCATCGCCCGAGC 1689

Db 3293 AGCCTATCTGGGTCTCAGGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 3333

```

RESULT 3
BX353556
LOCUS
DEFINITION
BX353556 Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED Homo sapiens
cDNA clone CS0DC008YE05 5-PRIME, mRNA sequence.
ACCESSION
BX353556
VERSION
BX353556.2 GI:46290133
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
REFERENCE
1. (bases 1 to 1016)
AUTHORS
Li W.B., Gruber C., Jessee J. and Polayes D.
TITLE
Full-length cDNA libraries and normalization
JOURNAL
Unpublished (2001)
COMMENT
On May 5, 2003 this sequence version replaced gi:30371778.
Contact: Genoscope
Genoscope - Centre National de Sequencage
2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen. This sequence belongs to sequence cluster
5694.f
For more information about this cluster, see
http://www.genoscope.cns.fr/cdna?s=CS0DC008AC030P1&c=5694.f.
FEATURES
source
1..1016
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DC008YE05"
/tissue_type="NEUROBLASTOMA"
/clone_lib="Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and EcoR V sites of the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."
ORIGIN
Query Match 54.9%; Score 928.6; DB 5; Length 1016;
Best Local Similarity 96.2%; Pred. No. 8.3e-198;
Matches 982; Conservative 19; Mismatches 14; Indels 6; Gaps 5;
QY 208 CGCCACGGCTGATGGTCTCTCTGGTGGCGTTGAGTCCACACGAAGGTCGAGATGCC 267
DB 1 CGCCACGGCTGATGGTCTCTCTGGTGGCGTTGAGTCCACACGAAGGTCGAGATGCC 60
QY 268 TGTCTGGGATGACACCGGCTGACGACTTCTGTCAAGGACCAAGGACGGCTAGCAGG 327
DB 61 TGTCTGGGATGACACCGGCTGACGACTTCTGTCAAGGACCAAGGACGGCTA-CAGG 119
QY 328 AGAACACCATGATCTTGACACGCTCTCCCGACACGGGTGAGCTGTGGCGGC 387
DB 120 AGAACACCATGATCTTGACACGCTCTCCCGACACGGGTGAGCTGTGGCGGC 179
QY 388 CTACGAGCAGTGGCGGAGCGGCGGACAGCGGGCTGTGGGACTACTCCCTGCACGT 447
DB 180 CTACGAGCAGTGGCGGAGCGGCGGACAGCGGGCTGTGGGACTACTCCCTGCACGT 239
QY 448 GGCATCACCGATGGATGAGAGCATCAAGAGGAGCTGGAGGCTCTGTGCAAGGAGAA 507
DB 240 GGCATCACCGATGGATGAGAGCATCAAGAGGAGCTGGAGGCTCTGTGCAAGGAGAA 299
QY 508 GGGTGAATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 567
DB 300 GGGTGAATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 359
QY 568 CCAGATGTACGAGATCTTACGATCATCCGGGACCTGGGGGCTTGGCCCGAGGTGCACGC 627

```

```

DB 360 CCAGATGTACGAGATCTTTCAGCATCATCCGGACCTGGGGGCTTGG -CCAGTGCACGC 418
QY 628 TGAGAACGGGGGACATCGTGGAGAGGAGCAGAAAGCGGTTGCTGGAGCTCGGCATCACTGG 687
DB 419 TGAGAACGGGGGACATCGTGGAGAGGAGCAGAAAGCGGTTGCTGGAGCTCGGCATCACTGG 478
QY 688 CCCCAGAGGCGCCACGTGCTCAGCCACCCAGAGAGGTGAGGGCTGAGGCGGTTGACCGAGC 747
DB 479 CCCCAGAGGCGCCACGTGCTCAGCCACCCAGAGAGGTGAGGGCTGAGGCGGTTGACCGAGC 538
QY 748 TGTACACATCGCCAAAGCAGGCAAACTGCCCCCTGTACCTGTCACCAAGGTGATGAGCAAGGG 807
DB 539 TGTACACATCGCCAAAGCAGGCAAACTGCCCCCTGTACCTGTCACCTGAGAGCAAGGG 598
QY 808 GGGGCGGAGCGCATCGCTCAGCCCAAGCGCAGAGGGGTGCTGCTGTTTGGGGAGGCCAT 867
DB 599 GGGGCGGAGCGCATCGCTCAGCCCAAGCGCAGAGGGGTGCTGCTGTTTGGGGAGGCCAT 658
QY 868 CACCGCCAGCTGGGGCAGCGACGGTTTCACTACTTGAGCAAGAACTGGGCGCAAGGCTGC 927
DB 659 CACCGCCAGCTGGGGCAGCGACGGTTTCACTACTTGAGCAAGAACTGGGCGCAAGGCGCG 718
QY 928 AGCTTTGCTCAGATCACTACCCCTGTCAACCCAGACCCACACGCGCAGACCACTCACTG 987
DB 719 AGCTTTGCTCAGATCACTACCCCTGTCAACCCAGACCCACACGCGCAGACCACTCACTG 778
QY 988 CTGCTGCTCAGGGGGGACCTCAGGTGACAGGCGGCGGCTGACCTTCACTGACCTGC 1047
DB 779 CTGCTGCTCAGGGGGGACCTCAGGTGACAGGCGGCGGCTGACCTTCACTGACCTGC 838
QY 1048 CCAGAAAGGCTGTGGGCAAGGACAACTTTCGCGCTGATCCCGAGGGCACCACCGGCAATTGA 1107
DB 839 CCAGAAAGGCTGTGGGCAAGGACAACTTTCGCGCTGATCCCGAGGGCACCACCGGCAATTGA 898
QY 1108 GGAGGCGCATGTCGAT -GGTCTGGAGAAATGTGTGCGCTCTGGAGAGATGGACGAGAATG 1166
DB 899 GGAGGCGCATGTCGATGGTGTGGGCAAGGACAACTTTCGCGCTGATCCCGAGGGCACCACCGGCAATTG 957
QY 1167 AGTTGCTGCGGTGACCATGACAAATGCTGCGCAAAATCTTCAATTTTACCAGGAAGG 1226
DB 958 ARTTCGTGCGGTGACMAGTMAA--TGCTGCAAAATTTTAAATTTTTCACCAAGRAGGG 1015
QY 1227 G 1227
DB 1016 G 1016

```

```

RESULT 4
BX353556
LOCUS
DEFINITION
BX353556 Homo sapiens
5', mRNA sequence.
AGENCOURT_8804382 NIH_MGC_47 Homo sapiens cDNA clone IMAGE:6377895
BX353556
VERSION
BX353556.1 GI:22372399
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1. (bases 1 to 938)
AUTHORS
NIH-MGC http://imgc.nhl.nih.gov/.
TITLE
National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL
Unpublished (1999)
COMMENT
Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Rubin Laboratory
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLCM2562 row: g column: 16

```

FEATURES		High quality sequence stop: 669. Location/Qualifiers 1. 938 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /clone="IMAGE:6377895" /tissue_type="neuroblastoma, cell line" /lab_host="DH10B (phage-resistant)" /clone_lib="NIH_MGC_47" /note="Organ: brain; Vector: pOTB7; Site 1: XhoI; Site 2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC Library."	
source		52.0%; Score 879; DB 5; Length 938; Query Match Best Local Similarity 98.4%; Pred. No. 1.1e-186; Matches 918; Conservative 0; Mismatches 12; Indels 3; Gaps 3;	
ORIGIN		<p>614 GCCAGGTGACGCTGAGAACGGGGACATCGTGGAGGAGGACGAAGCGGTTGCTGGAG 673</p> <p>1 GGCCAGGTGACGCTGAGAACGGGGACATCGTGGAGGAGGACGAAGCGGTTGCTGGAG 60</p> <p>674 CTGGCATCATCTGCCCGGAGGCGACGTCTCAGCCACCCCGAGGAGTGCGGCTGAG 733</p> <p>61 CTGGCATCATCTGCCCGGAGGCGACGTCTCAGCCACCCCGAGGAGTGCGGCTGAG 120</p> <p>734 GCGGTGATCCAGGCTGTACCATCGCCAGCAGGCAAACTGCCCGCTGTACGTCACCAAG 793</p> <p>121 GCGGTGATCCAGGCTGTACCATCGCCAGCAGGCAAACTGCCCGCTGTACGTCACCAAG 180</p> <p>794 GTGATGAGCAAGGGGGCGGCGACGCATCGTCCAGCCCAAGCGCAGAGGGGTGTCGTG 853</p> <p>181 GTGATGAGCAAGGGGGCGGCGACGCATCGTCCAGCCCAAGCGCAGAGGGGTGTCGTG 240</p> <p>854 TTTGGGAGGCCATCAGCGCGAGCTGGCCAGCCAGCGGTTCACTACTGAGCAAGAAC 913</p> <p>241 TTTGGGAGGCCATCAGCGCGAGCTGGCCAGCCAGCGGTTCACTACTGAGCAAGAAC 300</p> <p>914 TGGGCCAAGGCTGCAGGCTTGTGCATCATCCCCCTGTCAACCCAGACCCACACAGGCA 973</p> <p>301 TGGGCCAAGGCTGCAGGCTTGTGCATCATCCCCCTGTCAACCCAGACCCACACAGGCA 360</p> <p>974 GACCACCTCACTGCTTGTGTCCAGCGGGAGCCTCCAGGTGACAGGCGGCCCACTGC 1033</p> <p>361 GACCACCTCACTGCTTGTGTCCAGCGGGAGCCTCCAGGTGACAGGCGGCCCACTGC 420</p> <p>1034 ACCTTACCACTGCCAGAGGCTGTGGCAGAGCAAACTTCGGCTGTATCCCGAGGGC 1093</p> <p>421 ACCTTACCACTGCCAGAGGCTGTGGCAGAGCAAACTTCGGCTGTATCCCGAGGGC 480</p> <p>1094 ACCAAGGCTTGAAGGAGCGCATGTGCTGTGGGAGAAATGTGGGCTCTGGGAAG 1153</p> <p>481 ACCAAGGCTTGAAGGAGCGCATGTGCTGTGGGAGAAATGTGGGCTCTGGGAAG 540</p> <p>1154 ATGACGAGAAATGAGTTGCTCGCGGTGACCAAGTACAAATGCTGCAAAATCTTCAATTTT 1213</p> <p>541 ATGACGAGAAATGAGTTGCTCGCGGTGACCAAGTACAAATGCTGCAAAATCTTCAATTTT 600</p> <p>1214 TACCAAGGAGGGCGAGTGGTGTGGCTGTGACGCTGACCTGTGATATGAGACCCC 1273</p> <p>601 TACCAAGGAGGGCGAGTGGTGTGGCTGTGACGCTGACCTGTGATATGAGACCCC 660</p> <p>1274 AAGGCCACCAAGATCATCTCTGCCAAGACCCCAATCTGAACGTGGAGTACAACTTTC 1333</p> <p>661 AAGGCCACCAAGATCATCTCTGCCAAGACCCCAATCTGAACGTGGAGTACAACTTTC 720</p>	
FEATURES		<p>1. 1075 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /clone="IMAGE:5769897" /lab_host="DH10B" /clone_lib="NIH_MGC_121" /note="Organ: Brain; Vector: pCMV-SPORT6; Site 1: NotI; Site 2: EcoRV (destroyed); RNA source anonymous pool of 3 fetal brains, female age 20 weeks, female age 24 weeks, and male age 26 weeks. Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.7 kb, insert size range 0.7-3.5 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 017. Note: this is a NIH_MGC Library."</p>	
source		<p>1. 1075 Location/Qualifiers /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /clone="IMAGE:5769897" /lab_host="DH10B" /clone_lib="NIH_MGC_121" /note="Organ: Brain; Vector: pCMV-SPORT6; Site 1: NotI; Site 2: EcoRV (destroyed); RNA source anonymous pool of 3 fetal brains, female age 20 weeks, female age 24 weeks, and male age 26 weeks. Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.7 kb, insert size range 0.7-3.5 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 017. Note: this is a NIH_MGC Library."</p>	
ORIGIN		<p>Query Match Best Local Similarity 96.1%; Pred. No. 3.2e-185; Matches 938; Conservative 0; Mismatches 30; Indels 8; Gaps 4;</p> <p>1 GCGGCCCCCTACACAGAGACCCCGAGGAGGAGTGTCTTCCAGGCGCAAGAAAGCATCCC 60</p> <p>2 GCGGCCCCCTACACAGAGACCCCGAGGAGGAGTGTCTTCCAGGCGCAAGAAAGCATCCC 61</p> <p>61 CCGGATCAAGAGTACCGGCTTCTTGATCAGAGGTGGGAGGATCGTGAATGACGACCATC 120</p>	

FEATURES		High quality sequence stop: 669. Location/Qualifiers 1. 938 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /clone="IMAGE:6377895" /tissue_type="neuroblastoma, cell line" /lab_host="DH10B (phage-resistant)" /clone_lib="NIH_MGC_47" /note="Organ: brain; Vector: pOTB7; Site 1: XhoI; Site 2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC Library."	
source		52.0%; Score 879; DB 5; Length 938; Query Match Best Local Similarity 98.4%; Pred. No. 1.1e-186; Matches 918; Conservative 0; Mismatches 12; Indels 3; Gaps 3;	
ORIGIN		<p>614 GCCAGGTGACGCTGAGAACGGGGACATCGTGGAGGAGGACGAAGCGGTTGCTGGAG 673</p> <p>1 GGCCAGGTGACGCTGAGAACGGGGACATCGTGGAGGAGGACGAAGCGGTTGCTGGAG 60</p> <p>674 CTGGCATCATCTGCCCGGAGGCGACGTCTCAGCCACCCCGAGGAGTGCGGCTGAG 733</p> <p>61 CTGGCATCATCTGCCCGGAGGCGACGTCTCAGCCACCCCGAGGAGTGCGGCTGAG 120</p> <p>734 GCGGTGATCCAGGCTGTACCATCGCCAGCAGGCAAACTGCCCGCTGTACGTCACCAAG 793</p> <p>121 GCGGTGATCCAGGCTGTACCATCGCCAGCAGGCAAACTGCCCGCTGTACGTCACCAAG 180</p> <p>794 GTGATGAGCAAGGGGGCGGCGACGCATCGTCCAGCCCAAGCGCAGAGGGGTGTCGTG 853</p> <p>181 GTGATGAGCAAGGGGGCGGCGACGCATCGTCCAGCCCAAGCGCAGAGGGGTGTCGTG 240</p> <p>854 TTTGGGAGGCCATCAGCGCGAGCTGGCCAGCCAGCGGTTCACTACTGAGCAAGAAC 913</p> <p>241 TTTGGGAGGCCATCAGCGCGAGCTGGCCAGCCAGCGGTTCACTACTGAGCAAGAAC 300</p> <p>914 TGGGCCAAGGCTGCAGGCTTGTGCATCATCCCCCTGTCAACCCAGACCCACACAGGCA 973</p> <p>301 TGGGCCAAGGCTGCAGGCTTGTGCATCATCCCCCTGTCAACCCAGACCCACACAGGCA 360</p> <p>974 GACCACCTCACTGCTTGTGTCCAGCGGGAGCCTCCAGGTGACAGGCGGCCCACTGC 1033</p> <p>361 GACCACCTCACTGCTTGTGTCCAGCGGGAGCCTCCAGGTGACAGGCGGCCCACTGC 420</p> <p>1034 ACCTTACCACTGCCAGAGGCTGTGGCAGAGCAAACTTCGGCTGTATCCCGAGGGC 1093</p> <p>421 ACCTTACCACTGCCAGAGGCTGTGGCAGAGCAAACTTCGGCTGTATCCCGAGGGC 480</p> <p>1094 ACCAAGGCTTGAAGGAGCGCATGTGCTGTGGGAGAAATGTGGGCTCTGGGAAG 1153</p> <p>481 ACCAAGGCTTGAAGGAGCGCATGTGCTGTGGGAGAAATGTGGGCTCTGGGAAG 540</p> <p>1154 ATGACGAGAAATGAGTTGCTCGCGGTGACCAAGTACAAATGCTGCAAAATCTTCAATTTT 1213</p> <p>541 ATGACGAGAAATGAGTTGCTCGCGGTGACCAAGTACAAATGCTGCAAAATCTTCAATTTT 600</p> <p>1214 TACCAAGGAGGGCGAGTGGTGTGGCTGTGACGCTGACCTGTGATATGAGACCCC 1273</p> <p>601 TACCAAGGAGGGCGAGTGGTGTGGCTGTGACGCTGACCTGTGATATGAGACCCC 660</p> <p>1274 AAGGCCACCAAGATCATCTCTGCCAAGACCCCAATCTGAACGTGGAGTACAACTTTC 1333</p> <p>661 AAGGCCACCAAGATCATCTCTGCCAAGACCCCAATCTGAACGTGGAGTACAACTTTC 720</p>	
FEATURES		<p>1. 1075 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /clone="IMAGE:5769897" /lab_host="DH10B" /clone_lib="NIH_MGC_121" /note="Organ: Brain; Vector: pCMV-SPORT6; Site 1: NotI; Site 2: EcoRV (destroyed); RNA source anonymous pool of 3 fetal brains, female age 20 weeks, female age 24 weeks, and male age 26 weeks. Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.7 kb, insert size range 0.7-3.5 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 017. Note: this is a NIH_MGC Library."</p>	
source		<p>1. 1075 Location/Qualifiers /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /clone="IMAGE:5769897" /lab_host="DH10B" /clone_lib="NIH_MGC_121" /note="Organ: Brain; Vector: pCMV-SPORT6; Site 1: NotI; Site 2: EcoRV (destroyed); RNA source anonymous pool of 3 fetal brains, female age 20 weeks, female age 24 weeks, and male age 26 weeks. Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.7 kb, insert size range 0.7-3.5 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 017. Note: this is a NIH_MGC Library."</p>	
ORIGIN		<p>Query Match Best Local Similarity 96.1%; Pred. No. 3.2e-185; Matches 938; Conservative 0; Mismatches 30; Indels 8; Gaps 4;</p> <p>1 GCGGCCCCCTACACAGAGACCCCGAGGAGGAGTGTCTTCCAGGCGCAAGAAAGCATCCC 60</p> <p>2 GCGGCCCCCTACACAGAGACCCCGAGGAGGAGTGTCTTCCAGGCGCAAGAAAGCATCCC 61</p> <p>61 CCGGATCAAGAGTACCGGCTTCTTGATCAGAGGTGGGAGGATCGTGAATGACGACCATC 120</p>	

Db	62	CCGATCAGAGTGACCGCTTCTGATCAGAGTGGAGGATCGTGAATGACGACCGC	121	
Qy	121	CTTTTACGCTGATGTGACGCTGAAGATGCTTTGATAAAACAAATCGAGAAAACCTCAT	180	
Db	122	CTTTTACGCTGATGTGACGCTGAAGATGCTTTGATAAAACAAATCGAGAAAACCTCAT	181	
Qy	181	CGTCCCTGGGGCATCAAGACCAATGACGCCACGCGCTGATGGTCTTCTTGGTGGCGT	240	
Db	182	CGTCCCTGGGGCATCAAGACCAATGACGCCACGCGCTGATGGTCTTCTTGGTGGCGT	241	
Qy	241	TGACGCTCCACAAAGGCTGCAGATGCTGTCTGGGCATGACACCGGCTGACGACTTCTG	300	
Db	242	TGACGCTCCACAAAGGCTGCAGATGCTGTCTGGGCATGACACCGGCTGACGACTTCTG	301	
Qy	301	TCAGGGCACAAGGACGCTAGCAGAGGAACACCATGATCTTGGACCACTGTTTCCC	360	
Db	302	TCAGGGCACAAGGACGCTAGCAGAGGAACACCATGATCTTGGACCACTGTTTCCC	361	
Qy	361	CGACACGGGTGTAGGCTGTGCGGCGCTACGAGCAGTGGCGGAGCGGCGACAGCGC	420	
Db	362	CGACACGGGTGTAGGCTGTGCGGCGCTACGAGCAGTGGCGGAGCGGCGACAGCGC	421	
Qy	421	GGCTCTGCTGACTACTCCCTGCACGTGGACATCACCCGATGGCATGAGGATCAAGGA	480	
Db	422	GGCTCTGCTGACTACTCCCTGCACGTGGACATCACCCGATGGCATGAGGATCAAGGA	481	
Qy	481	GGAGCTGGAGGCCCTGTGACGAGAAAGGTGTGAATCTTCTTGGTCTTCAATGGCAT	540	
Db	482	GGAGCTGGAGGCCCTGTGACGAGAAAGGTGTGAATCTTCTTGGTCTTCAATGGCAT	541	
Qy	541	CAAGGACCGGTGACGTGACGACGACGACGACGACGACGACGACGACGACGACGACG	600	
Db	542	CAAGGACCGGTGACGTGACGACGACGACGACGACGACGACGACGACGACGACG	601	
Qy	601	CTTGGGGGCTTGGCCACAGGTGCACCTGAGAACGGGGACATCGTGGAGGAGGACGAA	660	
Db	602	CTTGGGGGCTTGGCCACAGGTGCACCTGAGAACGGGGACATCGTGGAGGAGGACGAA	661	
Qy	661	GGGGTGTGGAGCTGGCATCTGCGCCCGAGGGCCAGTGTCTGAGCCACCCGAGGA	720	
Db	662	GGGGTGTGGAGCTGGCATCTGCGCCCGAGGGCCAGTGTCTGAGCCACCCGAGGA	721	
Qy	721	GGTGGAGGCTGAGCGGTGTACCGAGCTGCACCATCGCCACGACGCAACCTGCCGCT	780	
Db	722	GGTGGAGGCTGAGCGGTGTACCGAGCTGCACCATCGCCACGACGCAACCTGCCGCT	781	
Qy	781	GTACGTCAACCAAGGTGATGAGCAAGGGGGCGCGCGCATCGCTCAGGCCAAGCGCAG	840	
Db	782	GTACGTCAACCAAGGTGATGAGCAAGGGGGCGCGCGCATCGCTCAGGCCAAGCGCAG	841	
Qy	841	A--GGGGTGTGTGTTTGGGAGCCCATCAGCCGACGCT--GGGACCGACGCTTCA	897	
Db	842	AAGGGTGTGTGTTTGGGAGCCCATCAGCCGACGCTTGGGGAAACCGACGCTTCA	901	
Qy	898	CTACTGGAGCAAGCAACTGGGCGCAA---GGCTGACGCTTCTGTCACATCA--CC	952	
Db	902	CTACTGGAGCAAGCAACTGGGCGCAAAGGGCGCGCGCTTCTGTCACATTAACCC	961	
Qy	953	AACCCAGACCCCAACCA 968		
Db	962	AACCAAGACCCCAACCA 977		
RESULT 6				
AK077000				
LOCUS				
DEFINITION				
AK077000				
VERSION				
KEYWORDS				

SOURCE	Mus musculus (house mouse)
ORGANISM	Mus musculus
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
AUTHORS	1 Carninci, P. and Hayashizaki, Y.
TITLE	High-efficiency full-length cDNA cloning
JOURNAL	Meth. Enzymol. 303, 19-44 (1999)
MEDLINE	99279253
PUBMED	10349636
REFERENCE	2
AUTHORS	Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
TITLE	Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes
JOURNAL	Genome Res. 10 (10), 1617-1630 (2000)
MEDLINE	20499374
PUBMED	11042159
REFERENCE	3
AUTHORS	Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Katsunai, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, S., Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
TITLE	RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer
JOURNAL	Genome Res. 10 (11), 1757-1771 (2000)
MEDLINE	20530913
PUBMED	11076861
REFERENCE	4
AUTHORS	The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium.
TITLE	Functional annotation of a full-length mouse cDNA collection
JOURNAL	Nature 409, 685-690 (2001)
REFERENCE	5
AUTHORS	The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.
TITLE	Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
JOURNAL	Nature 420, 563-573 (2002)
REFERENCE	6 (bases 1 to 1899)
AUTHORS	Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Katoh, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishii, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N., Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shingawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akashira, S., Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y.
TITLE	Direct Submission
JOURNAL	Submitted (16-APR-2002) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gs.riken.jp, URL: http://genome.gsc.riken.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)
COMMENT	cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. Please visit our web site for further details. URL: http://genome.gsc.riken.jp/ URL: http://phantom.gsc.riken.jp/ Location/Qualifiers
FEATURES	1. .1899
SOURCE	



Contact: Genoscope  
 Genoscope - Centre National de Sequencage  
 2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE  
 Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr  
 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime  
 end enriched, double-strand cDNA was digested with Not I and cloned  
 into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library  
 was not normalized. Library was constructed by Life Technologies, a  
 division of Invitrogen.  
 This sequence belongs to sequence cluster 5694.f  
 For more information about this cluster, see  
 http://www.genoscope.cns.fr/cdna?s=CS0DF015CD10QP1&c=5694.f.

## FEATURES

source

Location/Qualifiers  
 1..857  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="CS0DF015YH19"  
 /tissue\_type="FETAL BRAIN"  
 /dev\_stage="fetal"  
 /clone\_lib="Homo sapiens FETAL BRAIN"  
 /notes="Organ: brain; Vector: pCMVSPORT 6; 1st strand cDNA  
 was primed with a NotI-oligo(dT) primer. Five prime end  
 enriched, double-strand cDNA was digested with Not I and  
 cloned into the Not I and EcoRV sites of the pCMVSPORT 6  
 vector. Library was not normalized."

## ORIGIN

Query Match 48.7%; Score 823.2; DB 5; Length 857;  
 Best Local Similarity 99.4%; Pred. No. 3.9e-174;  
 Matches 855; Conservative 2; Mismatches 0; Indels 3; Gaps 3;  
 QY 5 CCCTACAGAGAGCCCGGAGGAGGATGCTTCAGGGGAGAAAGCATCCCCCG 64  
 DB 1 CCCTACAGAGAGCCCGGAGGAGGATGCTTCAGGGGAGAAAGCATCCCCCG 60  
 QY 65 ATCAGAGTGACCGCTTCTGATCAGAGGTGGAGGATCGTGAATGACGACCACTCTT 124  
 DB 61 ATCAGAGTGACCGCTTCTGATCAGAGGTGGAGGATCGTGAATGACGACCACTCTT 120  
 QY 125 TAGCTGATGTGACCTGGAAGATGCTTGATATAAAACAAATCGAGAAACCTCATCGTC 184  
 DB 121 TAGCTGATGTGACCTGGAAGATGCTTGATATAAAACAAATCGAGAAACCTCATCGTC 180  
 QY 185 CTTGGGGGATCAAGACCATGAGCCGACCGCTGATGCTCTTCTGTTGGCGTTGAC 244  
 DB 181 CTTGGGGGATCAAGACCATGAGCCGACCGCTGATGCTCTTCTGTTGGCGTTGAC 240  
 QY 245 GTCCACACAAGGCTGACAGATGCTGCTGGGATGACACCGGCTGACGACTTCTGTGAC 304  
 DB 241 GTCCACACAAGGCTGACAGATGCTGCTGGGATGACACCGGCTGACGACTTCTGTGAC 300  
 QY 305 GGCACCAAGCGAGCGCTAGCAGGAGGAACCAACATGATCTTGACCAACCGTCTTCCCGAC 364  
 DB 301 GGCACCAAGCGAGCGCTA-SAGGAGGAACCAACATGATCTTGACCAACCGTCTTCCCGAC 359  
 QY 365 ACGGTTGTAGCTGTGCGGCTTACGACAGTGGCGGAGCGGGGAGCAGCGCGCC 424  
 DB 360 ACGGTTGTAGCTGTGCGGCTTACGACAGTGGCGGAGCGGGGAGCAGCGCGCC 419  
 QY 425 TGTCTGCACTTACTCTTGCAGCTGGACATCACCAGTGGCATGAGAGCATCAAGGAGGAG 484  
 DB 420 TGTCTGCACTTACTCTTGCAGCTGGACATCACCAGTGGCATGAGAGCATCAAGGAGGAG 479  
 QY 485 CTGGAGGCGCTTGTCAAGGAGAGGGTGTGAATCTCTTCTGTTTTCATGGCATCAAG 544  
 DB 480 CTGGAGGCGCTTGTCAAGGAGAGGGTGTGAATCTCTTCTGTTTTCATGGCATCAAG 539  
 QY 545 GACCGTGTCCAGTGCAGCAGCAGCATGATGATGATCTTTCAGCATCATCCGGACCTG 604  
 DB 540 GACCGTGTCCAGTGCAGCAGCAGCATGATGATGATCTTTCAGCATCATCCGGACCTG 599  
 QY 605 GGGGCTTGGCCCGGAGTGCAGCTGAGAACCGGGAACATCTGTGGAGGAGGAGGAGCGG 664

DB 600 GGGGCTTGG-CCAGGTGCACGCTCAGAACGGGACATGTRGAGGAGGAGGAGCGG 658  
 QY 665 TTGCTGGAGCTCGGATCATCTGCCCCCGAGGCGCAGTGCTCAGCCACCCCGAGGAGGTG 724  
 DB 659 TTGCTGGAGCTCGGATCATCTGCCCCCGAGGCG-CACTGCTCAGCCACCCCGAGGAGGTG 717  
 QY 725 GAGGCTGAGGCGGTGTACCGAGCTGTCACTATGCGCCAGCAGGCAACCTGCCGCTGTAC 784  
 DB 718 GAGGCTGAGGCGGTGTACCGAGCTGTCACTATGCGCCAGCAGGCAACCTGCCGCTGTAC 777  
 QY 785 GTCAACCAAGCTGATGAGCAAGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 844  
 DB 778 GTCAACCAAGCTGATGAGCAAGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 837  
 QY 845 GTGCTGCTGTTTGGGAGGCC 864  
 DB 838 GTGCTGCTGTTTGGGAGGCC 857  
 RESULT 8  
 AK090197  
 LOCUS  
 DEFINITION  
 Mus musculus 11 days embryo spinal cord cDNA, RIKEN full-length  
 enriched library, clone:G630016E18 product: dihydropyrimidinase-like  
 4, full insert sequence.  
 ACCESSION  
 AK090197  
 VERSION  
 AK090197.1 GI:26105790  
 KEYWORDS  
 HTG; CAP trapper.  
 SOURCE  
 Mus musculus (house mouse)  
 ORGANISM  
 Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 REFERENCE  
 1  
 Carninci, P. and Hayashizaki, Y.  
 High-efficiency full-length cDNA cloning  
 Meth. Enzymol. 303, 19-44 (1999)  
 MEDLINE  
 99279253  
 PUBMED  
 10349636  
 2  
 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,  
 Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.  
 Normalization and subtraction of cap-trapper-selected cDNAs to  
 prepare full-length cDNA libraries for rapid discovery of new genes  
 Genome Res. 10 (10), 1617-1630 (2000)  
 MEDLINE  
 20493374  
 PUBMED  
 11042159  
 3  
 Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,  
 Konno, H., Akiyama, J., Nishi, K., Kitsuami, T., Tashiro, H., Itoh, M.,  
 Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,  
 Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,  
 Fujikawa, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, K.,  
 Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J.,  
 Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.  
 RIKEN integrated sequence analysis (RISA) system--384-format  
 sequencing pipeline with 384 multicapillary sequencer  
 Genome Res. 10 (11), 1757-1771 (2000)  
 MEDLINE  
 20530913  
 PUBMED  
 11076861  
 4  
 The RIKEN Genome Exploration Research Group Phase II Team and the  
 FANTOM Consortium.  
 Functional annotation of a full-length mouse cDNA collection  
 Nature 409, 695-690 (2001)  
 5  
 The FANTOM Consortium and the RIKEN Genome Exploration Research  
 Group Phase I & II Team.  
 Analysis of the mouse transcriptome based on functional annotation  
 of 60,770 full-length cDNAs  
 Nature 420, 563-573 (2002)  
 6  
 (bases 1 to 3161)  
 Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P.,



Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W., Hayashi, K., Hayatsu, N., Hiramoto, K., Hiroaka, T., Hirozane, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Kato, H., Kawah, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N., Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Takeda, Y., Tanaka, T., Tomaru, A., Toyota, T., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y.

Submitted (16-APR-2002) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration and Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: yoshihide-res@gsc.riken.jp, URL: <http://genome.gsc.riken.jp/>, Tel: 81-45-503-9222, Fax: 81-45-503-9216)

CDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.

URL: <http://genome.gsc.riken.jp/>  
URL: <http://fantom.gsc.riken.jp/>

5694.f  
For more information about this cluster, see  
http://www.genoscope.cng.fr/cdna?s=CS0DB005C06QP1&c=5694.f.

## FEATURES

source  
1. .963  
/location/Qualifiers  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="CS0DB005YF11"  
/issue\_type="NEUROBLASTOMA COT 10-NORMALIZED"  
/clone\_lib="Homo sapiens NEUROBLASTOMA COT 10-NORMALIZED"  
/notes="1st strand cDNA was primed with a NotI-oligo(dT)  
primer. Five prime end enriched, double-strand cDNA was  
digested with Not I and cloned into the Not I and EcoR V  
sites of the pCMVSPORT 6 vector. Library was normalized."

## ORIGIN

Query Match 47.6%; Score 804.6; DB 5; Length 963;  
Best Local Similarity 99.4%; Pred. No. 6e-170;  
Matches 848; Conservative 1; Mismatches 0; Indels 4; Gaps 4;

QY 12 CAGAGACCCCGAGGAGGATGCTCTCCAGGGCAAGAAACATCCCGGATCACGA 71  
DB 1 CAGAGACCCCGAGGAGGATGCTCTCCAGGGCAAGAAACATCCCGGATCACGA 60

QY 72 GTGACCGCTTCTGATCAGAGGTGGGAGGATCGTGAATGACGACCATGCTTTTACGCTG 131  
DB 61 GTGACCGCTTCTGATCAGAGGTGGGAGGATCGTGAATGACGACCATGCTTTTACGCTG 120

QY 132 ATGTGACCGTGGAGATGGCTTGATAAACAAATCGGAGAAACCTCATGTCCTCGGGG 191  
DB 121 ATGTGACCGTGGAGATGGCTTGATAAACAAATCGGAGAAACCTCATGTCCTCGGGG 180

QY 192 GCATCAAGACCATGACGCCACCGCTGATGTCTCTTCTGTGGGTGAGTGCACCA 251  
DB 181 GCATCAAGACCATGACGCCACCGCTGATGTCTCTTCTGTGGGTGAGTGCACCA 240

QY 252 CAAGGCTGCAGATGCTCTGCGGATGACACCGCTGACGACTTCTGTGAGGGACCA 311  
DB 241 CAAGGCTGCAGATGCTCTGCGGATGACACCGCTGACGACTTCTGTGAGGGACCA 300

QY 312 AGGACGCTAGCAGGAGGAACACCATGATCTTTGGACACGCTTCTCCCGACACGGGTG 371  
DB 301 AGGACGCTA-CAGGAGGAACACCATGATCTTTGGACACGCTTCTCCCGACACGGGTG 359

QY 372 TGAGCTGTGCGGGCTTACGACAGTGGCGGAGCGGGGAGACAGCGGCGCTGTGCG 431  
DB 360 TGAGCTGTGCGGGCTTACGACAGTGGCGGAGCGGGGAGACAGCGGCGCTGTGCG 419

QY 432 ACTACTCCCTGACGTTGACATCACCGGATGGCATGAGCATCAAGGAGGCTGGAGG 491  
DB 420 ACTACTCCCTGACGTTGACATCACCGGATGGCATGAGCATCAAGGAGGCTGGAGG 479

QY 492 CCTGGTCAAGGAGAGGGGTGAACCTCTCTCTGCTTTCATGGCATACAAGGACCGGT 551  
DB 480 CCTGGTCAAGGAGAGGGGTGAACCTCTCTCTGCTTTCATGGCATACAAGGACCGGT 539

QY 552 GCCAGTGCAGCGACGCGAGATGTACAGATCTTTCAGCATCATCCGGACCTGGGGCCT 611  
DB 540 GCCAGTGCAGCGACGCGAGATGTACAGATCTTTCAGCATCATCCGGACCTGGGGCCT 598

QY 612 TGGCCAGGTCACGCTGAGAACGGGACATCGTGGAGGAGGAGCAGAGCGGTTCGCTG 671  
DB 599 TGG-CCAGGTGCACGCTGAGAACGGGACATCGTGGAGGAGGAGCAGAGCGGTTCGCTG 657

QY 672 AGCTCGGCATCACTGGCCCCGAGGGCCAGTGTCTAGCCACCCCGAGGAGGTGAGGCTG 731  
DB 658 AGCTCGGCATCACTGGCCCCGAGGG-CACGTGTCTAGCCACCCCGAGGAGGTGAGGCTG 716

QY 732 AGGCGGTGTACCGAGCTGTACCATGCGCAGCAGGCAAACTGCCGCTGTAGTCAACCA 791  
DB 717 AGGCGGTGTACCGAGCTGTACCATGCGCAGCAGGCAAACTGCCGCTGTAGTCAACCA 776

QY 792 AGGTGATGAGCAAGGGGGCGCGCGCATCGCTCAGGCCAAGCGAGGGGTGGTGG 851  
DB 777 AGGTGATGAGCAAGGGGGCGCGCGCATCGCTCAGGCCAAGCGAGGGGTGGTGG 836

QY 852 TGTTTGGGAGCC 864  
DB 837 TGTTTGGGAGCC 849

RESULT 10  
BM928281 1105 bp mRNA linear EST 12-MAR-2002  
LOCUS AGENCOURT 6699776 NIH\_MGC\_121 Homo sapiens cDNA clone IMAGE:5770115  
DEFINITION 5', mRNA sequence.  
ACCESSION BM928281  
VERSION BM928281.1 GI:19378660  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 1105)  
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgaabs@mail.nih.gov  
Tissue Procurement: Life Technologies, Inc.  
cDNA Library Preparation: Life Technologies, Inc.  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
http://image.llnl.gov  
Plate: LLAM12832 row: k column: 12  
High quality sequence stop: 661.

FEATURES  
Location/Qualifiers  
1. .1105  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:5770115"  
/lab\_host="DH10B"  
/clone\_lib="NIH\_MGC\_121"  
/note="Organ: Brain; Vector: pCMV-SPORT6; Site 1: NotI;  
Site 2: EcoRV (destroyed); RNA source anonymous pool of 3  
fetal brains, female age 20 weeks, female age 24 weeks,  
and male age 26 weeks. Library is oligo-dT primed and  
directionally cloned (EcoRV site is destroyed upon  
cloning). Average insert size 1.7 kb, insert size range  
0.7-3.5 kb. Library is normalized and enriched for  
full-length clones and was constructed by C. Gruber  
(Invitrogen). Research Genetics tracking code 017. Note:  
this is a NIH\_MGC Library."

ORIGIN  
Query Match 47.1%; Score 795.8; DB 5; Length 1105;  
Best Local Similarity 93.9%; Pred. No. 5.8e-168;  
Matches 886; Conservative 0; Mismatches 43; Indels 15; Gaps 5;

QY 61 CCGGATCAGAGTGACCGCTTCTGATCAGAGGTGGGAGGATCGTGAATGACGACGATC 120  
DB 1 CCGGATCAGAGTGACCGCTTCTGATCAGAGGTGGGAGGATCGTGAATGACGACGATC 60

QY 121 CTTTACGCTGATGTGACGTGGAAGATGGCTTGATATAAACAATCGGAGAAAACCTCAT 180  
DB 61 CTTTACGCTGATGTGACGTGGAAGATGGCTTGATATAAACAATCGGAGAAAACCTCAT 120

QY 181 CTTCCCTGGGGGCATCAAGACCATTGAAGCCACCGCCCTGATGGTCTTCTCGGTGGCGT 240  
DB 121 CTTCCCTGGGGGCATCAAGACCATTGAAGCCACCGCCCTGATGGTCTTCTCGGTGGCGT 180

QY 241 TGACGTCCACACAGGCTGCAGATGCTCTCTGCGGCATGACACCGGCTGACGACACTTCTG 300

181 TGAGTGCACAAAGGCTGCAAGTCCCTGCTGCTGCGCATGACACGGCTGACGACTCTGTG 240  
301 TCAGGGACCAAGGACGAGCTAGCAGGAGGAGAACACCATGATCTTGGACCAAGCTTTCCC 360  
241 TCAGGGACCAAGGACGAGCTAGCAGGAGGAGAACACCATGATCTTGGACCAAGCTTTCCC 300  
361 CGACACGGGTGTGAGCTGCTGGCGGCTACGACAGTGGCGGAGCGGCGGACAGCGC 420  
301 CGACACGGGTGTGAGCTGCTGGCGGCTACGACAGTGGCGGAGCGGCGGACAGCGC 360  
421 GGCTGTGCGACTACTCCCTGCGAGTGGAGATCACCCGATGGCATGAGAGCATCAAGGA 480  
361 GGCTGTGCGACTACTCCCTGCGAGTGGAGATCACCCGATGGCATGAGAGCATCAAGGA 420  
481 GGACTGGAGGCCCTGTGCAAGGAGAGGGGTGCAACTCTCTCTGCTTTCATGGGCATA 540  
421 GGACTGGAGGCCCTGTGCAAGGAGAGGGGTGCAACTCTCTCTGCTTTCATGGGCATA 480  
541 CAAGGACCGGTGCGAGTGCAGCGACAGCCAGATGTACGAGATCTTCAAGCATCATCCGGGA 600  
481 CAAGGACCGGTGCGAGTGCAGCGACAGCCAGATGTACGAGATCTTCAAGCATCATCCGGGA 540  
601 CCTGGGGCCCTTGGCCCAAGTGCAGTGAACCGGGGACATCTGTGAGGAGGAGCAGAA 660  
541 CCTGGGGCCCTTGGCCCAAGTGCAGTGAACCGGGGACATCTGTGAGGAGGAGCAGAA 600  
661 GCGGTGCTGAGCTCGGCTACCTGCGCCCGAGGCGCCAGTGTCTCAGCCACCCCGAGGA 720  
601 GCGGTGCTGAGCTCGGCTACCTGCGCCCGAGGCGCCAGTGTCTCAGCCACCCCGAGGA 660  
721 GGTGAGGCTGAGGCGGTGTACCGAGCTGTACCATCGCCCAAGCAGGCAAACTGCCCGCT 780  
661 GGTGAGGCTGAGGCGGTGTACCGAGCTGTACCATCGCCCAAGCAGGCAAACTGCCCGCT 720  
781 GTAGCTACCAAGTGTAGTGAAGGGGGGCGCCGACGATCTCAGCCCAAGCGCAG 840  
721 GTAGCTACCAAGTGTAGTGAAGGGGGGCGCCGACGATCTCAGCCCAAGCGCAG 780  
841 AGGGGTGGTC--GTGTTTGGGGAGCCCATCACCGCCAGCTGGGCAACCGAC-----GGT 892  
781 AGGGGTGGTCCCGGTTTGGGGAGCCCATCACCGCCAGCTGGGCAACCGACCGGGTT 840  
893 TCACACTACTGG--AGCAAGAACTGGGCGCAAGG---CTGCAGCGCTTCTGTCACATCAACCCC 947  
841 CCCCCCTACTGGGAGGCAAGAACTGGGCGCCAGGCCCCCAGGCTTTCGGCACTCAACCCC 900  
948 C--TGTCACCCAGACCCCAAGCGGAGACCACTTCACTTGT 989  
901 CCTTGTCAACCCAAAGCCCGCCACCGGTGACCCCTNCCCTGGT 944

## RESULT 11

BQ068024

LOCUS

AGENCOURT 6792381 NIH\_MGC\_121 Homo sapiens cDNA clone IMAGE:5769039

DEFINITION

5', mRNA sequence.

ACCESSION

BQ068024

VERSION

EST.

KEYWORDS

SOURCE

ORGANISM

Homo sapiens (human)

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

1 (bases 1 to 1058)

NIH-MGC http://mgi.nci.nih.gov/

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)

Contact: Robert Strauberg, Ph.D.

Email: csaps-r@mail.nih.gov

DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
http://image.llnl.gov

Plate: LUAM12829 row: n column: 16

High quality sequence stop: 676.

## FEATURES

## source

1..1058  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/Clone="IMAGE:5769039"  
/lab\_host="DH10B"  
/clone\_lib="NIH\_MGC\_121"  
/notes="Organ: b'rain; Vector: pCMV-SPORT6; Site\_1: Notif;  
Site\_2: EcoRV (destroyed); RNA source anonymous pool of 3  
fetal brains, female age 20 weeks, female age 24 weeks,  
and male age 26 weeks. Library is oligo-dT primed and  
directionally cloned (EcoRV site is destroyed upon  
cloning). Average insert size 1.7 kb, insert size range  
0.7-3.5 kb. Library is normalized and enriched for  
full-length clones and was constructed by C. Gruber  
(Invitrogen). Research Genetics tracking code 017. Note:  
this is a NIH\_MGC Library."

## ORIGIN

Query Match 47.0%; Score 795; DB 5; Length 1058;  
Best Local Similarity 97.5%; Pred. No. 8.7e-168;  
Matches 818; Conservative 0; Mismatches 17; Indels 4; Gaps 1;  
QY 1 GC CGC C C C C T A C C A G A G A C C C C C A G G A G C A G A T G C T T C C C A G G C C A A G A A A G C A T C C C 60  
DB 69 GC CGC C C C T A C C A G A G A C C C C C A G G A G C A G A T G C T T C C C A G G C C A A G A A A G C A T C C C 128  
QY 61 C C G G A T C A C G A G T G A C G C C T T C T G A T C A G A G T G G G A G G A T C G T G A A T G A C G A C C A G T C 120  
DB 129 C C G G A T C A C G A G T G A C G C C T T C T G A T C A G A G T G G G A G G A T C G T G A A T G A C G A C C A G T C 188  
QY 121 C T T T T A C G C T G A T G C A C G T G G A A G A T G C T T T G A T A A A C A A A T C G G A G A A A C C T C A T 180  
DB 189 C T T T T A C G C T G A T G C A C G T G G A A G A T G C T T T G A T A A A C A A A T C G G A G A A A C C T C A T 248  
QY 181 C G T C C C T T G G G G C A T C A A G A C C A T T G A C C G C C A C G C C T G A T G C T C C T C T G T G C G C T 240  
DB 249 C G T C C C T T G G G G C A T C A A G A C C A T T G A C C G C C A C G C C T G A T G C T C C T C T G T G C G C T 308  
QY 241 T G A C G T C C A C A C A A G G C T G C A G A T G C T G T C T T G G G C A T G A C A C C G C T G A C C A C T T C T G 300  
DB 309 T G A C G T C C A C A C A A G G C T G C A G A T G C T G T C T T G G G C A T G A C A C C G C T G A C C A C T T C T G 368  
QY 301 T C A G G G C A C C A A G G C A G C C T A G C A G A G A A C C A C A T G A T C T T G G A C C A C C G T C T T C C C 360  
DB 369 T C A G G G C A C C A A G G C A G C C T A G C A G A G A A C C A C A T G A T C T T G G A C C A C C G T C T T C C C 428  
QY 361 C G A C A C G G T G T G A G C T G C T G C G C C C T A C A G A G A G T G C G G A G C A G C G G C G A C A G C G C 420  
DB 429 C G A C A C G G T G T G A G C T G C T G C G C C C T A C A G A G A G T G C G G A G C A G C G G C G A C A G C G C 488  
QY 421 G G C T G C T G C G A C T A C T C C C T G C A C G T G G A C A T C A C C C G A T G C A T G A G A G A T C A A G G A 480  
DB 489 G G C T G C T G C G A C T A C T C C C T G C A C G T G G A C A T C A C C C G A T G C A T G A G A G A T C A A G G A 548  
QY 481 G G A G C T G A G G C C C T G G T C A A G A G A A G G G T G A A C T C C T C C T G T C T T C A T T G G C A T A 540  
DB 549 G G A G C T G A G G C C C T G G T C A A G A G A A G G G T G T G A A C T C C T C C T G T C T T C A T T G G C A T A 608  
QY 541 C A A G G A C C G G T G C A G T G C A G C A G C A G C A G A T G T A C G A G A T C T T C A G C A T C A T C C G G G A 600  
DB 609 C A A G G A C C G G T G C A G T G C A G C A G C A G C A G A T G T A C G A G A T C T T C A G C A T C A T C C G G G A 668  
QY 601 C T T G G G G C C T T T G G C C C A G T G C A C C T C A G A A C C G G G A C A T C G T G A G A G A G A G A G A G A A 660  
DB 669 C T T G G G G C C T T T G G C C C A G T G C A C C T C A G A A C C G G G A C A T C G T G A G A G A G A G A G A G A A 728



Email: seqref@genoscope.cns.fr, Web : [www.genoscope.cns.fr](http://www.genoscope.cns.fr)  
 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime  
 end enriched, double-strand cDNA was digested with Not I and cloned  
 into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library  
 was not normalized. Library was constructed by Life Technologies, a  
 division of Invitrogen.

This sequence belongs to sequence cluster 5694.f

For more information about this cluster, see  
<http://www.genoscope.cns.fr/cdna?s=CS0DF028CF05Q0P1&c=5694.f>.

## FEATURES

source

```
1. 1042
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DF028YL09"
/tissue_type="FETAL BRAIN"
/dev_stage="fetal"
/clone_lib="Homo sapiens FETAL BRAIN"
/notes="Organ: brain; Vector: pCMVSPORT_6; 1st strand cDNA
was primed with a NotI-oligo(dT) primer. Five prime end
enriched, double-strand cDNA was digested with Not I and
cloned into the Not I and EcoRV sites of the pCMVSPORT 6
vector. Library was not normalized."
```

## ORIGIN

```
Query Match 46.7%; Score 790; DB 5; Length 1042;
Best Local Similarity 91.0%; Pred. No. 1.2e-166;
Matches 829; Conservative 24; Mismatches 56; Indels 2; Gaps 2;

QY 1 GCGCCCTACAGAGACCCCGAGGAGGAGGATGCTTCCAGGGCGAAGAAAGCATCCC 60
DB 113 KCGCCCTACAGAGACCCCGAGGAGGAGGATGCTTCCAGGGCGAAGAAAGCATCCC 172
QY 61 CCGGATCAGGAGTACCGCTTCTGATCAGAGTGGGAGGATCGTGAATCAGCAGATC 120
DB 173 CCGGATCAGGAGTACCGCTTCTGATCAGAGTGGGAGGATCGTGAATCAGCAGATC 232
QY 121 CTTTACGCTGATGTGCACTGGAAGATGGCTTGATATAAAACAAATCGAGAAACCTCAT 180
DB 233 CTTTACGCTGATGTGCACTGGAAGATGGCTTGATATAAAACAAATCGAGAAACCTCAT 292
QY 181 CGTCCCTGGGGGATCAAGACCATGACCGCCAGCTGATGCTTCTCGTGGCGGT 240
DB 293 CGTCCCTGGGGGATCAAGACCATGACCGCCAGCTGATGCTTCTCGTGGCGGT 352
QY 241 TGACCTCCACAAAGGCTGAGATGCTGCTGGGATGACACCGGCTGACGACTTCG 300
DB 353 TGACCTCCACAAAGGCTGAGATGCTGCTGGGATGACACCGGCTGACGACTTCG 412
QY 301 TCAGGGCACCAGGCGCTAGCAGGAGGAACCAACCATGATCTTGACCAACGCTTTCCC 360
DB 413 TCAGGGCACCAGGCGCTAGCAGGAGGAACCAACCATGATCTTGACCAACGCTTTCCC 472
QY 361 CGACAGGGGTGAGCTTGGCGGCTTACGAGAGTGGCGGAGCGGCGGAGCAGCGC 420
DB 473 CGACAGGGGTGAGCTTGGCGGCTTACGAGAGTGGCGGAGCGGCGGAGCAGCGC 532
QY 421 GGCTGCTGAGCTACTCTCCCTGACGAGGAGGATCACCCTGATGATGATGATGATGAT 480
DB 533 GGCTGCTGAGCTACTCTCCCTGACGAGGAGGATCACCCTGATGATGATGATGATGAT 592
QY 481 GGAGCTGGAGCCCTGCTCAAGGAGAGGGTGTGAATCTCTCTGCTTTCATGGCATA 540
DB 593 GGAGCTGGAGCCCTGCTCAAGGAGAGGGTGTGAATCTCTCTGCTTTCATGGCATA 652
QY 541 CAAGGACCGGTGCGAGTGCAGCGACGACCGAGATGTACGAGATTTTCAGCATATCCGGGA 600
DB 653 CAAGGACCGGTGCGAGTGCAGCGACGACCGAGATGTACGAGATTTTCAGCATATCCGGGA 712
QY 601 CCTGGGGGCTTGGCCAGGTGCACTGAGAGCGGGAGATCGTGGAGGAGGAGCAGAA 660
DB 713 CCTGGGGGCTTGG-CCAGGTGCACTGAGAGCGGGAGATCGTGGAGGAGGAGCAGAA 771
```

```
QY 661 GCGTTTCTGGAGCTCGGCATCACTGCCCCCGAGGGCCAC-GTGCTCAGCCACCCCGAGG 719
DB 772 GCGTTTCTGGAGCTCGGCATCACTGCCCCCGAGGGSCACAGTGTCTCAGCCACCCCGAGG 831
QY 720 AGGTGAGGCTGAGGCGGTATCCGAGCTGTACACCTCCGAGCAGGCAAACTGCCCGC 779
DB 832 AGGTGAGGCTGAGGCGGTATCCGAGCTGTACACCTCCGAGCAGGSAATCTGCCCGC 891
QY 780 TGTACCTTACCAGGTGTAGCAGCAAGGGGGCGGCGCCAGCCATCGCTCAGGCCCAAGCGCA 839
DB 892 TGTACCTTACCAGGTGTAGCAGCAAGGGGGCGGCGCCATCCTCARGCCCAACCCAAA 951
QY 840 GAGGGGTGTGTTTGGGAGCCCATCACCGCCAGCTTGGCCACCGAGCTTTCACACT 899
DB 952 GGGKTTCTGTGTTGGRRCCATCACCCCAACCTGGSACCCGCTTACACATWATKASAAA 1011
QY 900 ACTGGAGCAAG 910
DB 1012 AMTKGCMAG 1022
```

## RESULT 14

BX444582

LOCUS

DEFINITION

BX444582

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

On May 15, 2003 this

Contact: Genoscope

Genoscope - Centre National de Sequencage

2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE

Email: seqref@genoscope.cns.fr, Web : [www.genoscope.cns.fr](http://www.genoscope.cns.fr)

1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime

end enriched, double-strand cDNA was digested with Not I and cloned

into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library

was not normalized. Library was constructed by Life Technologies, a

division of Invitrogen.

This sequence belongs to sequence cluster 5694.f

For more information about this cluster, see

<http://www.genoscope.cns.fr/cdna?s=CS0DN004AF01Q1P1&c=5694.f>.

Location/Qualifiers

1. 900

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="CS0DN004YK01"

/tissue\_type="ADULT BRAIN"

/dev\_stage="adult"

/clone\_lib="Homo sapiens ADULT BRAIN"

/notes="Organ: brain; Vector: pCMVSPORT\_6; 1st strand cDNA

was primed with a NotI-oligo(dT) primer. Five prime end

enriched, double-strand cDNA was digested with Not I and

cloned into the Not I and EcoRV sites of the pCMVSPORT 6

vector. Library was not normalized."

## ORIGIN

Query Match 45.9%; Score 775.4; DB 5; Length 900;

Best Local Similarity 98.3%; Pred. No. 2.2e-163;

Matches 794; Conservative 9; Mismatches 3; Indels 2; Gaps 2;

QY 1 GCGCCCTTACAGAGACCCCGAGGAGGAGGATGCTTCCAGGGCGAAGAAAGCATCCC 60

DB 95 GCGCCCTTACAGAGACCCCGAGGAGGAGGATGCTTCCAGGGCGAAGAAAGCATCCC 154

QY	61	CCGGATCAGAGTGACCGCTTCTGATCAGAGTGGGAGATCGTGAATGACGACCAAGTC	120	JOURNAL	Science 302 (5652), 1960-1963 (2003)
Db	155	CCGGATCAGAGTGACCGCTTCTGATCAGAGTGGGAGATCGTGAATGADNWBAGTC	214	PUBMED	14671302
QY	121	CTTTTACGCTGATGTGACAGTGAAGATGGCTTTGATAAAACAAATCGGAGAAACCTCAT	180	REFERENCE	2 (bases 1 to 1680)
Db	215	CTTTTACGCTGATGTGACAGTGAAGATGGCTTTGATAAAACAAATCGGAGAAACCTCAT	274	AUTHORS	Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarawal,A., Todd,M.A., Tanenbaum,D.M., Civeillo,D.R., Lu,F., Murphy,B., Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M.
QY	181	CGTCCCTGGGGGATCAAGACATTGACCGCCACGGCCTGATGGTCTTCTGGTGGCGT	240	TITLE	Direct Submission
Db	275	CGTCCCTGGGGGATCAAGACATTGACCGCCACGGCCTGATGGTCTTCTGGTGGCGT	334	JOURNAL	Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive, Rockville, MD 20850, USA
QY	241	TGACGTCCACACAAAGCTGCAGATGCTGTCTTGGGCATGACACCGCTGACGACTTCG	300	COMMENT	These sequences were made by sequencing genomic exons and ordering them based on alignment.
Db	335	TGACGTCCACACAAAGCTGCAGATGCTGTCTTGGGCATGACACCGCTGACGACTTCG	394	FEATURES	Location/Qualifiers
QY	301	TCAGGGCACCAAGCGAGCGCTAGCAGAGGAAACACCATGATCTTGGACCACTTTCCC	360	source	1..1680
Db	395	TCAGGGCACCAAGCGAGCGCTAGCAGAGGAAACACCATGATCTTGGACCACTTTCCC	454	organism	"/organism="Mus musculus"
QY	361	CGACAGGGTGTAGCTGTGCGGCTTACGAGCATGTGGCGGAGCGGCGGACAGCGC	420	molecule	"/mol_type="genomic DNA"
Db	455	CGACAGGGTGTAGCTGTGCGGCTTACGAGCATGTGGCGGAGCGGCGGACAGCGC	514	db_xref	"/db_xref="taxon:10090"
QY	421	GGCTGCTGGACTACTCCCTGCACCTGGACATCACCCGATGGCATGAGAGCATCAAGGA	480	gene	<1..>1680
Db	515	GGCTGCTGGACTACTCCCTGCACCTGGACATCACCCGATGGCATGAGAGCATCAAGGA	574	gene	"/gene="CRMP1"
QY	481	GGAGCTGGAGGCCCTGGTCAAGAGAGGGTGTGAATCTCTTCTGGTCTTTCATGGCATA	540	locus_tag	"/locus_tag="HOM6794"
Db	575	GGAGCTGGAGGCCCTGGTCAAGAGAGGGTGTGAATCTCTTCTGGTCTTTCATGGCATA	634	ORIGIN	
QY	541	CAAGGACCGGTGCAGTGCAGCAGCAGCCAGATGTACGAGATCTTCAGCATCATCCGGGA	600	Query Match	45.7%; Score 772.4; DB 9; Length 1680;
Db	635	CAAGGACCGGTGCAGTGCAGCAGCAGCCAGATGTACGAGATCTTCAGCATCATCCGGGA	694	Best Local Similarity	63.5%; Pred. No. 1.1e-162;
QY	601	CTTGGGGGCTTGGCCAGCGTGCACCTGAGACGGGACATCGTGGAGGAGGAGCAGAA	660	Matches 1028; Conservative	0; Mismatches 591; Indels 0; Gaps 0;
Db	695	CTTGGGGGCTTGG-CAAGGTGCACCTGAGAACGGGACATCGTGGAGGAGGAGCAGAA	753	QY	71 AGTGACCGCTTCTGATCAGAGTGGAGGATCGTGAATGACGACGATCCTTTACGCT
QY	661	GCGGTGTCTGGAGCTCGGCATCACTGGCCCCGAGGGCCAGTGTCTCAGCCACCCCGAGGA	720	Db	1 AGTGACCGCTTCTGATCAGAGTGGAGGATCGTGAATGACGACGATCCTTTACGCT
Db	754	GCGGTGTCTGGAGCTCGGCATCACTGGCCCCGAGGG-CAAGTGTCTCAGCCACCCCGAGGA	812	QY	131 GATGTGCAGTGGAGATGGCTTTGATAAAACAAATCGGAGAAACCTCATCGTCCCTGGG
QY	721	GGTGGAGGCTGAGCGGTGTACCGAGCTGTACCATCGCCATCGCAGCAGCAGAACTGCCCGCT	780	Db	61 GATGTCTACTAGAGATGGACTCTATAAAACAAATAGGAGAGAACCTGATTTCTCTGGT
Db	813	GGTGGAGGCTGAGCGGTGTACCGAGCTGTACCATCGCCAGCAGGCAAACTGCHCGST	872	QY	191 GGCATCAAGACCATTTGACGCCACCGCTGATGGTCTTCTCTGGTGGCGTTGACGTCCAC
QY	781	GTACGTCAACAAGGTGATGACGAGGGG	808	Db	121 GGAGTGAAGACCATCGAGCGGAATGGCCGAAATGGTCAATTCGCCGTGGCATTTGATGTCAAC
Db	873	GTACGTCAACAAGGTGATGACGAGGGG	900	QY	251 ACAGGCTGCAGATGCTTCTGGGCATGACACCGCTGACGACTTCTGTGAGGACACC
RESULT 15	AY419150	1680 bp	DNA	linear	GSS 17-DEC-2003
LOCUS	AY419150	Genomic survey sequence.			
DEFINITION	AY419150	Mus musculus CRMP1 gene, VIRTUAL TRANSCRIPT, partial sequence.			
ACCESSION	AY419150				
VERSION	AY419150.1	GI:39775110			
KEYWORDS	GSS.				
SOURCE	Mus musculus (house mouse)				
ORGANISM	Mus musculus				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.				
AUTHORS	1 (bases 1 to 1680)				
	Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarawal,A., Todd,M.A., Tanenbaum,D.M., Civeillo,D.R., Lu,F., Murphy,B., Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M.				
TITLE	Inferring nonneutral evolution from human-chimp-mouse orthologous gene trios				





**THIS PAGE BLANK (USPTO)**

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 23, 2005, 23:11:38 ; Search time 1102 Seconds  
(without alignments)  
2509.354 Million cell updates/sec

Title: US-09-367-496C-7

Perfect score: 1690

Sequence: 1 gcgcgccctaccagagacc.....atgaccacatcgccgcgacgc 1690

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents NA:\*

1: /cgn2\_6/ptodata/1/ina/5A COMB.seq.\*

2: /cgn2\_6/ptodata/1/ina/5B COMB.seq.\*

3: /cgn2\_6/ptodata/1/ina/6A COMB.seq.\*

4: /cgn2\_6/ptodata/1/ina/6B COMB.seq.\*

5: /cgn2\_6/ptodata/1/ina/PCTUS COMB.seq.\*

6: /cgn2\_6/ptodata/1/ina/backfiles1.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	965.4	57.1	4403	4	US-09-949-016-5632
2	963.8	57.0	5421	4	US-09-814-915A-66
3	956.6	56.6	1719	4	US-09-949-016-199
4	876.8	51.9	2841	4	US-09-949-016-4735
5	875.2	51.8	2842	4	US-09-702-705-1800
6	875.2	51.8	2842	4	US-09-736-457-1800
7	875.2	51.8	2842	4	US-09-671-325-1800
8	850.6	50.3	5032	4	US-09-949-016-4321
9	850.6	50.3	5047	4	US-09-949-016-200
10	821.8	50.9	3331	4	US-09-976-594-131
11	389.2	23.0	1524	4	US-09-252-991A-6970
12	389.2	23.0	1611	4	US-09-252-991A-6924
13	389.2	23.0	1782	4	US-09-252-991A-6769
14	389.2	23.0	1909	4	US-09-949-016-5326
15	363.8	21.5	1488	4	US-09-734-237B-61
16	317.6	18.8	2113	4	US-10-114-810-1
17	295.4	17.5	1491	4	US-09-734-237B-63
18	229.6	13.6	2100	3	US-09-836-992-2
19	189.2	11.2	1374	4	US-08-415-658-20
20	188	11.1	1632	4	US-09-270-767-1366
21	188	11.1	1632	4	US-09-270-767-16648
22	130.8	7.7	1383	1	US-08-289-709-2
23	130.8	7.7	1383	1	US-08-602-656-2
24	130.4	7.7	1380	4	US-09-950-772-3
25	130.4	7.7	3343	4	US-09-950-772-7
26	116.4	6.9	336	4	US-09-270-767-1866
27	116.4	6.9	336	4	US-09-270-767-17148

28	112	6.6	601	4	US-09-949-016-22041	Sequence 22041, A	
29	112	6.6	601	4	US-09-949-016-199950	Sequence 199950, A	
30	110.8	6.6	601	4	US-09-949-016-22040	Sequence 22040, A	
31	110.8	6.6	601	4	US-09-949-016-199949	Sequence 199949, A	
32	110.8	6.6	81433	4	US-09-949-016-11941	Sequence 11941, A	
33	110.8	6.6	84227	4	US-09-949-016-17374	Sequence 17374, A	
C	34	109.6	6.5	601	4	US-09-949-016-168120	Sequence 168120, A
35	109.6	6.5	72549	4	US-09-949-016-16477	Sequence 16477, A	
36	106	6.3	601	4	US-09-949-016-22078	Sequence 22078, A	
37	106	6.3	601	4	US-09-949-016-200054	Sequence 200054, A	
38	102.2	6.0	66988	4	US-09-949-016-11942	Sequence 11942, A	
39	102.2	6.0	66989	4	US-09-949-016-16063	Sequence 16063, A	
C	40	96.6	5.7	601	4	US-09-949-016-22109	Sequence 22109, A
C	41	96.6	5.7	601	4	US-09-949-016-154109	Sequence 154109, A
42	91.4	5.4	444	4	US-09-513-999C-8209	Sequence 8209, Ap	
43	87.8	5.2	601	4	US-09-949-016-22077	Sequence 22077, A	
44	87.8	5.2	601	4	US-09-949-016-200053	Sequence 200053, A	
45	81.8	4.8	1377	4	US-09-916-501A-4	Sequence 4, Appli	

#### ALIGNMENTS

#### RESULT 1

US-09-949-016-5632  
; Sequence 5632, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 5632  
; LENGTH: 4403  
; TYPE: DNA  
; ORGANISM: Human  
US-09-949-016-5632

Query Match	57.1%	Score	965.4	DB	4	Length	4403
Best Local Similarity	73.6%	Pred. No.	4.1e-215				
Matches	1230	Conservative	0	Mismatches	441	Indels	0
Gaps	0						
QY	20	CCAGAGAGAGATGTCCTTCAGGCGAAGAAAGCATCCCGGATCAGGATGACCGC	79				
DB	310	CCAGAGAGAGATGTCCTTCAGGCGAAGAAAGCATCCCGGATCAGGATGACCGC	369				
QY	80	CTTCTGATCAGAGTGGAGGATCGTAATGACGACAGTCCTTTTACGCTGATGTCAC	139				
DB	370	CTTCTGATCAGAGGAGTAAATTTGTAATGATGACGATGCTTATGACGATATAC	429				
QY	140	GTGGAGATGGCTTGATAAAACAAATCGGAGAAACCTCATCGTCCCTGGGGCATCAAG	199				
DB	430	ATGGAGATGGCTTGATCAAGCAATAGGAGAAATCTGATTCGACGAGGAGTGAAG	489				
QY	200	ACCATTCAGCGCCACCGGCTGATGTCCTTCGTTGGGGTTCAGCTCCACACAGGCTG	259				
DB	490	ACCATTCAGCGCCACCTCCCGGATGTCCTTCGTTGGGGTTCAGCTCCACACAGGCTG	549				
QY	260	CAGATGCTGCTGCGGATGACACCGGCTGACGCTTCTGTGAGGCGGACCAAGGCGG	319				
DB	550	CAGATGCTGCTGCGGATGACGCTTCTGTGAGGCGGACCAAGGCGGCGCC	609				
QY	320	CTAGCAGGAGGAACCAACCATGATCTTGGACCATGCTTCCCCACACGCGGTGTGAGCCTG	379				

```
Db 610 CTGGCTGGGGAAACCACTATGATCATTTGACCACTGTTCTTGAGCTGGGACAAGCCTG 669
Qy 380 CTGGCGGCTTACAGCAGTGGCGGGAGCGGGCGGACAGCGCGCTGCTGCGACTACTCC 439
Db 670 CTGGCTGCTTTGACCACTGGAGGGAATGGCGCGACAGCAAGCTCTGCTGAGACTACTCT 729
Qy 440 CTGCACTGTGACATCAACCCGATGGCATGAGAGCATCAAGAGAGAGCTGGAGGCGCTGGTC 499
Db 730 CTGCATGTGGACATCAAGCAGTGGCATTAAGGGCATCCAGGAGAGATGGAAAGCGCTTGTG 789
Qy 500 AAGGAGAAAGGTGTGAACCTCTTCTGCTGCTTCAATGSCATACAAGAGACCGGTGCCAGTGC 559
Db 790 AAGGATCAAGGGGTAAATCTCTCTGCTGCTATGCTGCTTCAAGATGCGCTTCAGACTA 849
Qy 560 AGCGACAGCCAGATGTACGAGATCTTTCAGCATCATCCGGACCTGGGGGCTTTGGGCCAG 619
Db 850 ACGGATTCGACATTTATGAAGTACTGAGTGTGATCCGGGATATTTGGCGCCATAGCCCAA 909
Qy 620 GTGCAGCTGAGAACCGGGACATCGTGGAGGAGGACGAGCGGTTGCTGGAGCTCGGC 679
Db 910 GTCACGCGAGAAATGGGCACATCATTTGCGAGGAGCAGCAGAGGATCTTGGATCTGGGC 969
Qy 680 ATCACTGGGCCCCGAGGCCACGTGCTCAGCCACCCGAGAGGTGGAGGCTGAGGGCGTG 739
Db 970 ATCAAGGCCCCGAGGACATGTGCTGAGCGGACCTGAGAGGTGCGAGCCGAAGCCGTG 1029
Qy 740 TACGAGCTGTCAACATGCGCAAGCAGGCAAACTGCGCGCTGTACGTCACCAAGGATG 799
Db 1030 AATCGTGCCATCAACATCGCAACACAGACCAACTGCGCGCTGTATATCAACAAGGTGATG 1089
Qy 800 AGCAAGGGGCGCGCAGCGCATCGCTCAGGCCAAGCGCAGAGGGGTGCTGTTTGGG 859
Db 1090 AGCAAAAGCTCTGCTGAGGTGCTGCGCCAGGCGACGGAAGGGAACCTGTTGGTATGCG 1149
Qy 860 GAGCCCATCAACCGGACCTGGGCGACCGAGCGTTTCACTACTGAGCAAGAACTGGGCC 919
Db 1150 GAGCCCATCACTGCCAGCTTGGGACGAGCGGTCTCCATCTACTGAGCAAGAACTGGGCC 1209
Qy 920 AAGGCTGAGCCTTGTGATCATCAACCCCTGTCAACCCAGACCCCAACCGGACCAAC 979
Db 1210 AAGGCTGCTGCTTGTGATCACTCCCACTCCCACTGAGCGCTGTATCAACCACTCCAGACTTT 1269
Qy 980 CTCACTGCTGCTGCTGAGCGGAGCTTCAAGTGAAGGAGCGGCGGCTGACCTTC 1039
Db 1270 CTCACTGCTGCTGCTGAGAGCTTCCAGGTCAAGGTCAGCGGAGTGGCCATTCGACGTTT 1329
Qy 1040 ACCACTGCCAGAGGCTGTGGGCAAGGACAACTTGGCGCTGATCCCGGAGGCGACCAAC 1099
Db 1330 AACACTGCCAGAGGCTGTAGGAAAGGACAACTTCACTGATTCGGAGGGGACCAAT 1389
Qy 1100 GGCAATTGAGAGCGCATGTGATGCTTGGGAAATGTGTGCGCTTGGGAAAGATGGAC 1159
Db 1390 GGCACCTGAGGAGCGGATGTCGCTCATCTGGGCAAGGCTGTGGTCACTGGGAAAGATGGAT 1449
Qy 1160 GAGAAATGCTTCTGCGGCTGACAGTACAAATGCTGCGCAAAATCTCAATTTTACCCA 1219
Db 1450 GAGAAACGATTTGTGGCTGTGACAGGACCAATGCGAGCAAAAGTCTTCAACCTTTACCCC 1509
Qy 1220 AGGAAAGGGCGAGTGGCTGTGGGCTCTGAGCGCTGACCTGTGTCATATGGAAACCCCAAGGCC 1279
Db 1510 CGGAAAGGCCGATGCTGTGGGATCCGATGCGGACCTGTGTCATCTGGGACCCCGACAGC 1569
Qy 1280 ACCAAGATCATCTGCGAAGACCCCAATCTGAAAGTGGAGTACAAATCTTTCGAGGGA 1339
Db 1570 GTTAAAAACCATCTCTGCAAGACACAAACAGCTCTCTCGAGTACAAATCTTGAAGGC 1629
Qy 1340 GTGAGTGGCGGGAGCGCTGCGGTGCTAATAGTCAGGGCGAGTGGCGCTGGAGGAC 1399
Db 1630 ATGAGTGGCGGCTCTCCCACTGTTGGTCAATCAGCCAGGGGAAGATTTGCTTGGAGGAC 1689
Qy 1400 GGGAAAGATGTTGTGACCCCGGGGCGGCGCTTCTGCTCCCTCGGAAAAAATTTCCCGGAC 1459
```

```
Db 1690 GGCACCCCTGATGTCTACCGAAGGCTCTGAGAGCTACTATTCCCGGAAGCCCTTCCCTGAT 1749
Qy 1460 TTTGTCTACAAGAGGATCAAAGCTCGCAACAGGCTGGCGGAGATCAACGGTGTGCCCGCT 1519
Db 1750 TTTGTCTACAAGGATCAAAGCTCGCAACAGGCTGGCGGAGATCAACGGTGTGCCCGCT 1809
Qy 1520 GGGCTGTATGACGGGCGCTTCCACAGGCTGATGTCCTGCGCAAGCCAGGAGTGGCGCT 1579
Db 1810 GGGCTGTATGACGGGCGCTTCCACAGGCTGATGTCCTGCGCAAGCCAGGAGTGGCGCT 1869
Qy 1580 CCGGCGCGGCGCTTCCACAGGCTGATGTCCTGCGCAAGCCAGGAGTGGCGCT 1639
Db 1870 TCCCTGGCGCAAGAGCTTCTCCGCAAGCAGGAGGCGGCGGAGTGGCGCT 1929
Qy 1640 TCGGGGTTTACGCTATCTGGGCTTCAAGGCTGATGACCAATCGCCCGAGCGC 1690
Db 1930 TCTGGATTCAAGTTGTTCTGCTGCTCAGATTGATGACAAATTTCCCGCGCGC 1980

RESULT 2
US-09-814-915A-66
; Sequence 66, Application US/09814915A
; Patent No. 6750015
; GENERAL INFORMATION:
; APPLICANT: Horwitz, Kathryn
; APPLICANT: Richer, Jennifer
; TITLE OF INVENTION: Progesterone Receptor-Regulated Gene Expression and Methods Relate
; TITLE OF INVENTION: Therefo
; FILE REFERENCE: 2848-39
; CURRENT APPLICATION NUMBER: US/09/814,915A
; CURRENT FILING DATE: 2002-03-21
; PRIOR APPLICATION NUMBER: 60/214,870
; PRIOR FILING DATE: 2000-06-28
; NUMBER OF SEQ ID NOS: 108
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 66
; LENGTH: 5421
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-814-915A-66

Query Match 57.0%; Score 963.8; DB 4; Length 5421;
Best Local Similarity 73.5%; Pred. No. 1e-214;
Matches 1229; Conservative 0; Mismatches 442; Indels 0; Gaps 0;

Qy 20 CCAGAGCAGGATGTCCTTCCAGGGCAAGAAAGCATCCCGGATCAGGATGACCGC 79
Db 1325 CCAGAGCAGGATGTCCTTATCAGGGCAAGAAATATTTCCAGCATCACGAGCGATCGT 1384
Qy 80 CTTCTGATCAGAGTGGGAGGATCGTGAATGACGACCGCTCTTTTACGCTGATGTCAC 139
Db 1385 CTTCTGATCAGAGGAGGTAAATTTGTAATGATGACGAGTCTTATGACAGACATATAC 1444
Qy 140 GTGGAAGATGCGCTTGTATATAAAACAAATCGGAGAAAACCTCATCGCTCCGCGGATCAAG 199
Db 1445 ATGGAAGATGCGCTTGTATCAAGCAATAGGAGAAATCTGATTGCGCAGGAGGTGAAG 1504
Qy 200 ACATTTGACGCCCAACCGGCTGATGTCCTTCTGCTGGGCGGTTGACGCTTCAACAGGCTG 259
Db 1505 ACCATCGAGGCCCACTCCCGGATGTCCTCCCGGAGGAAATTTGACGCTTCAACCTCGCTTC 1564
Qy 260 CAGATCGCTGCTCTGGGATGACACCGGCTGACGACTTCTGTCAGGCGCACCAAGGAGCG 319
Db 1565 CAGATCGCTGATCAGGGAATGACGCTGCTGCTGATGATTTCTTCCAGGAACAAGGCGCC 1624
Qy 320 CTAGCAGGAGGAACCAACCATGATCTTTGGACCACTCTTCCCGACACGCGGTGTGAGCTG 379
Db 1625 CTGGCTGGGGAACCACTATGATCATTTGACCACTGTTGTTCTGAGCTGGGACAGGCTG 1684
Qy 380 CTGGCGGCTTACGAGCAGTGGCGGAGCGGAGCAGCGCGGCTGCTCGGACTACTCC 439
Db 1685 CTGGCTGCTTTTACCAAGTGGAGGAAATGGGCGCGACAGCAAGTCTCTGCTGTGACTACTCT 1744
```



Qy	512	GTGAACTCCTTCTCGTGTCTTCATGGCATA	CAAGGACCGGTGGCCAGTGCAGCGCAGCAG	571	
Db	481	GTAATTTCTTCTCGTGTACATGGCTTTCAA	AGATCGCTTCCAGCTTAA	CGGATTTGCCAG 540	
Qy	572	ATGTACGAGATCTTCACGATCATCCGGGAC	CTCGGGGCTTTGGCCACAGTGCACGCTGAG	631	
Db	541	ATTTATGAAGTACTGAGTGTGATTCGGGAT	ATTGGCCCATAGCCCAAGTCCACGCGAA	600	
Qy	632	AA	CGGGACATCTGTGGAGGAGGACAGAAGCGGT	TGCTGGAGCTCGGCATCACTTGGCCCC 691	
Db	601	AATGGCAGCATATTCAGAGGAGCAGAGGAT	CTCTGATCTGGGCATCA	CGGGCCCC 660	
Qy	692	GAGGGC	ACGTGTCTCAGCCACCCCGAGGAGTGGAGCT	GAGCGGTGTACCGAGCTGT 751	
Db	661	GAGGGA	CAATGTCTGAGCGCACCTGAGGAGGT	TCGAGGCGGAAGCCGTGAATCTGTGCCCATC 720	
Qy	752	ACCATCGCCAAAGCAGGCAAACTGCCCGCT	GTACGTACCAAGTGTATGACGAAGGGGGCG	811	
Db	721	ACCATCGCCAA	CCGAGCCAACTGCCCGCTGTATATCA	CCAAGGTGATGAGCAAAAGCTCT 780	
Qy	812	GCGAGCGCCATCGCTCAGGCGCAAGCGCAG	AGGGGTGTCTGTGTTTGGGAGGCCATCA	CACC 871	
Db	781	GCTGAGTGCATCGCCGAGGACCGGAAGAG	GGAACCTGTGTGTATGCGGAGGCCCATCACT	840	
Qy	872	GCCAGCCTGGGCACCGACGGTTCA	CACTACTGGAGCAGAACTCTGGGCCAAGGCT	TCGACGC 931	
Db	841	GCCAGCTTGGGA	CGGACGGCTCCCAATCTGGAGCAAGAACT	GGGCCAAGGCTGCTGCC 900	
Qy	932	TTGCTC	ACATCA	CCCCCTGTCAACCCAGACCCCAACGAGGGGAGACCACTCA	CACTGCTTG 991
Db	901	TTTGTCA	CCCTCCCA	CCCTTGAGCCCTGATCCAA	CCCACTCCAGACTTCTCAACTCCCTTG 960
Qy	992	CTGTCCAGCGGGACCTCCAGGTG	CAGGACAGGCGCCCACTGCACCTTCA	CCACTTCCGCCAG 1051	
Db	961	CTGTCTGTGAGACCTCCAGGT	CACGGCAGTGCCCATTGCA	CGTTTAACTGTCCCCAG 1020	
Qy	1052	AAGGCTGTGGCCAAAGGACAACTTCG	CGCTGATCCCGAGGGCACCA	CGGCATTTGAGGAG 1111	
Db	1021	AAGGCTGTAGGA	AGGACAACTTCA	CCCTGATTCGGAGGGCACCAATGGCACTGAGGAG 1080	
Qy	1112	CGCATGTGATGTCTGGAGAAATGTGTGG	CCCTCTCGGAGAGATGGA	CGAGATGAGTTTC 1171	
Db	1081	CGGATGTCCGTCACTCGGCA	AGGCTGTGTGTCACTGGGAAGATGGA	TGAGAACCACTTTT 1140	
Qy	1172	GTCCGGGTGACAGGTACAAATGTCTGC	GCCTCTCGGAGAGATGGA	CGAGATGAGTTTC 1231	
Db	1141	GTGGCTGTGAC	CAGCAATATGCAGCCAAAGTCTTCAA	CCCTTTTACCCCGGGAAGGCCGC 1200	
Qy	1232	GTGGCTGTGGCTCTGACGCTGACCTCGT	CATATGGAACCCCAAGGCCACA	CAAGATCATC 1291	
Db	1201	ATTGCTGTGGATCCGATGCCGACTGTGT	CATCTGGACCCCGACAGGTTTAA	ACCATC 1260	
Qy	1292	TCTGCCAAGACCCCAATCTGAA	CGTGGAGTACAACATCTTTCGAGGAGT	TGGAGTGC CGG 1351	
Db	1261	TCTGCCAAGACA	CAACAGCTCTCTCGAGTACA	ACATCTTTGAAGCATGAGTGC CGC 1320	
Qy	1352	GGAGCGCTGCCGTGTGTCATTAAGT	CAGGCCGAGTGGCGCTGGAGGACGGGA	GATGTTT 1411	
Db	1321	GGCTCCCCA	CTGTGTGTTC	CAGCCAGGGGAAGATTGTCTGGAGGACGGCACCCCTGCAT 1380	
Qy	1412	GTCA	CCCGGGGGCGGCGCTCGTCCCTCGGA	AAACATTTCCCGGACTTTGTCTACAAG 1471	
Db	1381	GTCA	CCGAAAGGCTGTGACGCTACA	TTTCCCGGAAGCCCTTCCCTGATTTTGTTTACAG 1440	
Qy	1472	AGGATCAAAGCTCGCA	CAGGCTCGCGGAGATCA	CACGGTGTGCCCTCGTGCGCTGTATGAC 1531	
Db	1441	CGTATCAAGG	CAAGCAGCGCTGGCTGAGCTGAG	AGGGTTCTCTGTGGCGCTGTATGAC 1500	
Qy	1532	GGGCGCTC	ACAGAGTGTGTGCTTCC	CAAGCAGGAGTGGCGCTCGGCCCGCGCG 1591	
Db	1501	GGA	CTGTGTGTAAGTGTCTGTGAC	CGCCCAAGACATCTCAGCGCTCTCTGCGCCAG 1560	
Qy	1592	TCCTGCC	CAGCAAGATCTCGTGCCTCTCTGTG	CGCAACCTTACATCACTGCGGGTTCAGC 1651	

## RESULT 4

US-09-949-016-4735  
; Sequence 4735, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 4735  
; LENGTH: 2841  
; TYPE: DNA  
; ORGANISM: Human  
US-09-949-016-4735

Query Match	51.9%	Score 875.8;	DB 4;	Length 2841;
Best Local Similarity	70.0%	Pred. No. 1.6e-194;		
Matches 1181; Conservative	0;	Mismatches 507;	Indels	0;
Gaps	0;			

Qy	3	CGCCCCATCAGAGACCCCCAGAGAGCAGGATGCTCTTCCAGGGGCGAAGAAAGCATCCCCC	62
Db	122	CGGCCGGCCGAGGGGCCAGAGGGGGCCATGTCGTACCAAGGGCGAAGAGAGCATCCCGC	181
Qy	63	GGATCAGAGTGACCGCGCTTCTGATCAGAGGTGGGAGGATCGTGAATGACGACCAAGTCCT	122
Db	182	ACATCAGAGTGACCGGACTCTCATCAAAGGTGACGGATCATCAACGATGACCAATCCC	241
Qy	123	TTTATGCTGATGTGCACGTGGAGATGGCTTGATATAAAACAAATCGGAGAAAACCTCATCG	182
Db	242	TTTATGCTGAGCTCTACTCGGAGATGGACTTATCAAAACAAATAGGAGAGAACTTAATCG	301
Qy	183	TCCCTGGGGGCATCAAGACCAATTGACGCCACGGCCCTGATGGTCTTCTTCGGTGGCGTTG	242
Db	302	TTCTCTGGTGGATGAGACCAATTGAGCCAAACGGGCGGATGGTTATTTCGGAGGATTG	361
Qy	243	ACGTCCACACAAGGCTCGAGATGCTGTCTCTGGGCATGACACCGGCTGACGACTTCTGTGC	302
Db	362	ATGTCAACACGTACTCTGACGAAGCCCTCCACAGGGGATGACTGCGGCTGATGACTTCTTCC	421
Qy	303	AGGGCACCAAGGCAGCGCTAGCAGGAGGAGACCACTATGATCTTGGACCAAGTCTTCCCG	362
Db	422	AAGGGACAGGGCGGCACTGGTGGCGGGACCAAGATGATCATTTGACCATGTTGTTCTCTG	481
Qy	363	ACAGGGTGTGAGCCTCTGTCGGCCCTACAGACGATGGCGGGAGCGGGCGGACGACGCGG	422
Db	482	AACCTGGTCCAGCCTACTGACCTCTTTTCGAGAAAGTGGCAACGAAGCAGCTGACACCAAT	541
Qy	423	CCTGCTGGACTACTCCCTGCAAGTGAGACATCAACCGATGGCATGAGGACATCAAGGAGG	482
Db	542	CCTCTCTGTGATTAATCCCTCTCAAGTGAGACATCAACAGCTGGTATCGATGGCGGCTCGGAGG	601
Qy	483	AGCTGGAGGCCCTGGTCAAGGAGAAAGGGTGTGAATCTCTTCTCGGTCTTTCATGGCATACA	542
Db	602	AGCTGGAGGGTCTGGTGCAGCAAAAGGCGTCAATTCTTCCAAAGTCTACATGGCCCTATA	661



Db 602 AGCTGGAGTGTCTGTGTGACGACAAAGGCGCTCAATTCTCCAAAGTCTACATGCGCCTATA 661  
Qy 543 AGGACCGGTGCTGAGTGCAGGACGACGAGATGTACGAGATCTTCAGCATATCCGGGACC 602  
Db 662 AGGATGTCTACCAAAATGTGCGACGACGAGCTCTATGAAGCCTTTTACCTCTTTAAGGGCC 721  
Qy 603 TGGGGGCTTTGGGCCAGAGTGCACGCTGAGAACGGGACATCGTGGAGGAGGACAGAAAGC 662  
Db 722 TGGGAGCTGTGATCTTGGTCCATGCGAGAAATGGAGATTTGTAGCTCAGGAACAAAGC 781  
Qy 663 GGTGTGCTGAGAGCTCGGATCACTTGGCCCCGAGGCGACGCTGTGCTGACGACCCCGAGAGG 722  
Db 782 GATCTCTGAGATGGGCATCAGCGGTCCCGAGGCGCATGCTCTGGCCAGGAGACCTGAAAGC 841  
Qy 723 TGGAGCTGAGGGGTGTACCGAGCTGTACCAATCGCAAGCAGGCGAACTGCGCGCTGT 782  
Db 842 TGGAGCGCGAGGGGTGTTCGGGGCCATCAACATTCGGGCGGATCAACTGCGCCTGTGT 901  
Qy 783 ACCTCACCAGGTGATGAGCAAGGGGCGGCGGACGACATCGCTCAGGCGCAAGCGCAGAG 842  
Db 902 ACATCACCAGGTCTATGAGCAAGGTGACGCGGACATCATCGCTCTGGCCAGGAGAAAG 961  
Qy 843 GGTGTGTGTGTTGGGAGCCCATCAACGCGCAGCTTGGGACCGAGCGTTTCACACTACT 902  
Db 962 GGGCCCTAGTTTTTGGAGAGCCCATTTGCGGCGACGCTTGGGACCGATGGCACCATTAAT 1021  
Qy 903 GGAGCAAGACTGGCCCAAGGCTGACGCTTGTCTGTCAGCGGGGACCTCCAGGTGACAGGCA 1022  
Db 1022 GGAGCAAGAACTGGCCCAAGGCTGCGGCGTTCGTGACTTCCCTCCCTCGAGCCCGGACC 1081  
Qy 963 CCACCAAGGAGACCACTCAGCTGTGTGTCAGCGGGGACCTCCAGGTGACAGGCA 1022  
Db 1082 CTACCAAGCGGACTTGTGACTTCTCTACTTGGCTGTGGGACTTGCAGTCAAGGCA 1141  
Qy 1023 GCGCCCATGTGCACTTCAACCTGCGCAGAGGCTGTGGGCAAGGCAAACTTCGCGCTGA 1082  
Db 1142 GCGGCCACTGTCTCCTACAGCACTGCGCCAGAGGCGGTGGGCAAGGCAAACTTTACCTGA 1201  
Qy 1083 TCCCGAGGGCACCAAGGCAATTTGGAGGCGATGTCGATGCTGGGAGAAATGTGTGG 1142  
Db 1202 TCCCGAGGGGTGCAACGGGATAGAGGCGGATGACCGGTGCTTGGGCAAGGCGGTGG 1261  
Qy 1143 CCTCTGGAGAGTGGACGAGATGAGTTCTGTCGGGTGACAGTCAAAATGCTGCCAAA 1202  
Db 1262 CTACTGGCAAAATGGATGAGAACGAGTTTGTGCTGTCTACCAAGCAACATGAGGACCA 1321  
Qy 1203 TCTTCAATTTTATCCCAAGAGGCGGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1262  
Db 1322 TCTTTAACTGTATCCCAAGGAAAGGCGGATTTGCGGTGGCTCGGATGCGGACGTGTGTCA 1381  
Qy 1263 TATGGAAACCCCAAGGCGCAAGATCATCTCTGCGAAGACCCCAATCTGAACTGTGAGT 1322  
Db 1382 TCTGGGACCCCGCAAGGTTTGAAGACCAATAACAGCCAAAGTCAAAAGTCTGCGGCTGAGT 1441  
Qy 1323 ACAACATCTTCAGGAGTGTGAGTGTGCGGAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1382  
Db 1442 ACAACATCTTCAGGAGTGTGAGTGTGCGGAGCGCTGCTGCTGCTGCTGCTGCTGCTG 1501  
Qy 1383 GAGTGGCGCTGGAGGACGGAAGATGTTTGTGTCAACCGGGGCGGCGCTTGTGTGTGTGTGT 1442  
Db 1502 AGATCGTCTTTGAAGACGGAACATCAACGCTCAACAGGCGATGGGCGGCTTCAATTCGCG 1561  
Qy 1443 GGAACATCTCCGAGCTTGTCTACAGAGGATCAAGCTCGCAACAGGCTGCGGAGGAG 1502  
Db 1562 GGAAGCGCTTCCGAGGACCTGTATACAGCGCTCAAAATCAAGAAATGAAGTTTGTGGAT 1621  
Qy 1503 TCCAGGTGTGCGCGGTGTATGAGCGGCGCTGCTGAGGAGTGTGCTGCTGCTGCTGCTGCTG 1562  
Db 1622 TGAAGGGGTTTCAAGGGGCACTGTATGAGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1681  
Qy 1563 AGCAGGAGTGGGCTTCCGGGCGGCGGTCTGCTGCGGAGGAGATCTCGCTGCTGCTGCTG 1622  
Db 1682 AATATGCAACTCCCGCTCTTTCAGCCAAATCTTTCGCTTCTTAAACACAGCCGCCCA 1741

Qy 1623 TGGCAACCTACATCATGCTGGGTTTCCAGCTATCTGGGTTCTCAGGCTGATGACCAATCG 1682  
Db 1742 TCGAAACCTCCACCACTTCAAGCTTATCAGCTTATCAGTGTCCCGAGATGATGACAAATC 1801  
Qy 1683 CCCGAGC 1690  
Db 1802 CCAGGCGC 1809

## RESULT 6

US-09-736-457-1800  
; Sequence 1800, Application US/09736457  
; Patent No. 6509448  
; GENERAL INFORMATION:  
; APPLICANT: Wang, Tongtong  
; APPLICANT: Bangur, Chaitanya S.  
; APPLICANT: Lodes, Michael A.  
; APPLICANT: Fanger, Gary  
; APPLICANT: Vedvik, Tom  
; APPLICANT: Carter, Darrick  
; APPLICANT: Retter, Marc  
; APPLICANT: Mannion, Jane  
; APPLICANT: Fan, Liqun  
; APPLICANT: Wang, Aijun  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND  
; FILE REFERENCE: 210121.478C15  
; CURRENT APPLICATION NUMBER: US/09/736,457  
; CURRENT FILING DATE: 2000-12-13  
; NUMBER OF SEQ ID NOS: 1864  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 1800  
; LENGTH: 2842  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-736-457-1800

Query Match 51.8%; Score 875.2; DB 4; Length 2842;  
Best Local Similarity 69.9%; Pred. No. 3.7e-194;  
Matches 1180; Conservative 0; Mismatches 508; Indels 0; Gaps 0;

Qy 3 CGCCCCCTACACAGACCCCGAGGACGAGTGTCTTCCAGGGCAAGAAAAGCATCCCC 62  
Db 122 CGCCCGGCGAGGGCGCCAGAGGGGCCATGTGTACAGGGCAAGAGAGCATCCGC 181  
Qy 63 GGATCACGAGTGACCGCTTCTGTATCAGAGGTGGGAGGATCGTGAATGACGACAGTCT 122  
Db 182 ACATCACGAGTGACCGACTCTCTCATCAAAGGTGGACGGATCATCAACGATGACCAATCC 241  
Qy 123 TTATCGCTGATGTGACGTGGAGATGGCTTGTGATATAAAACAAATCGGAGAAAACCTCATCG 182  
Db 242 TTATGCTGACGTCTACCTGGAGGATGGACTTATCAAAACAAATAGGAGAACTTAATCG 301  
Qy 183 TCCTCGGGGCGATCAAGACCATTTGACGCCACCGCTGTGCTTCTTCTGGTGGGCTTG 242  
Db 302 TTCTGCTGGATGAAGACCATTTGAAGCCAAACGGGCGGATGTATTATCCGGAGGTATTG 361  
Qy 243 AGCTCCACAAAGGCTGACAGATGCTGTCTGTGGGATGACACCGCTGACGACTTCTGTC 302  
Db 362 ATGTCAACACGTACTGTGCAAGAGCCCTCCAGGGGATGACTGCGGCTGATGACTTCTTCC 421  
Qy 303 AGGGCAACGAGCGCTAGCAGGAGAACCAACATGATCTTGGACCAAGCTTCTCCCG 362  
Db 422 AAGGGACCAAGGGCGGACCTGGTGGGCGGACCAACGATGATCATTTGACCATGTTGTTCTG 481  
Qy 363 ACAACGGTGTGAGCTGTGCGGCGCTAGCAGCAGTGGGCGGAGCGGCGGAGCAGCGCG 422  
Db 482 AACCTGGTCCAGCTTACTGACCTCTTTCGAGAGTGGCCAGCAGCTGACACCAAT 541  
Qy 423 CTGTGTGAGTACTTCTCCCTGACGCTGGACATCAACCGATGGGATGAGAGATCAAGAGG 482  
Db 542 CTGTGTGATTAATCTCCCTCCAGTGGACATCAACAGCTGGTACGATGGCGTTTCGGAGG 601





```
Db 542 CCTGCTGTGTAATTAACCTCCCTCCACGTCGACATCAAAAGCTGGTACGATGGGTTTCGGGAGG 601
Qy 483 AGCTGAGGCGCCCTGGTCAAGGAGAAAGGTGTGAACTCTCTTCCTGGTGTCTTCAATGGCAATACA 542
Db 602 AGCTGAGGCTGTGGTGCAGGACAAAGGCGTCAATTCCTTCCAAAGTCTATACATGGCCCTATA 651
Qy 543 AGGACCGGTGCAGTGCAGGACGAGCAGACAGATGTACGAGATCTTTCAGCATCATCCGGGACC 602
Db 662 AGATGTCTTACCAATATGTCCGACAGCCAGCTCTATGAAGCCTTTACCTCTCTTAAAGGCC 721
Qy 603 TGGGGGCTTTGGCCCAAGGTGCACGCTGAGAACGGGACATCGTGGAGGAGGACAGAAAGC 662
Db 722 TGGGAGCTGTGATCTTGGTCCATGTCAGAAAATGAGATTGTAGCTCAGGAACAAAAGC 781
Qy 663 GGTTCGTGGAGCTCGGACATCACTGGCCCCCGAGGGCCACAGTGTCTAGCCACCCCGAGGAGG 722
Db 782 GGATCCTGGAGATGGGCATCACGGGTCCCGAGGGCCATGCCCTGAGCAGACCTTGAAGAGC 841
Qy 723 TGGAGGCTGAGGCGGTGTACCGAGCTGTACCAATCGCAAGCAGGCAAACTGCCCGCTGT 782
Db 842 TGGAGGCGGAGGCGGTGTTCGGGCCATCAACCAATGGGGCCCGATCAACTGCCCCCTGTGT 901
Qy 783 ACCTCAACCAAGGTGATGAGCAAGGGGGCGGCGGACGCCATCGCTCAGGCCCAAGCGCAGAG 842
Db 902 ACATCAACCAAGGTGATGAGCAAGGTGACGCGGACGACATCATCGCTCTGGCCAGGAAGAAG 961
Qy 843 GGTGTGTGTGTTTGGGAGGCCATCAACCGCCAGCCTTGGGACCGGACGATGTTCACTACT 902
Db 962 GGCCCCCTAGTTTTTGGAGAGCCCAATTCGCCAGACCTTGGGGACCGATGGCACCCCATTTACT 1021
Qy 903 GGAGCAAGAACTGGGCGCCAGGCTGCGAGCCTTTCGTCAATCACCCCTGTCAACCCAGACC 962
Db 1022 GGAGCAAGAACTGGGCGCCAGGCTGCGGCGGTTCGTGACTTCCCTCCCTGAGCCCGGACC 1081
Qy 963 CCACCAAGGACGACCACTCACCTGCTTGTCTCAGCGGGGACCTCCAGGTGACAGGCA 1022
Db 1082 CTACCAAGCGCGACTACTTGACCTCCCTACTTGGCGCTGTGGGACTTTCAGGTTCACAGGCA 1141
Qy 1023 GCGCCCACTGCACTTTCACCACTGCCCCAGAAAGGCTGTGGGCAAGGACAACTTCGCGCTGA 1082
Db 1142 GCGGCCACTGTCTCCTACAGCACTGCCCCAGAAAGGCGTGGGCAAGGACAACTTTACCCCTGA 1201
Qy 1083 TCCCCGAGGCGACCAAGCGGATTTGAGGAGCGCATGTCGATGCTCTGGGAAATGTGTGG 1142
Db 1202 TCCCCGAGGCGTCAACCGGATAGAGAGCGGATGACCGCTGCTGTGGGACAGGCGGTGG 1261
Qy 1143 CCTCTGGGAAGATGGACGAGAATGAGTTTCGTGCGGCTGACCAAGTACAAATGTCGCCAAAA 1202
Db 1262 CTACTGGCAAAATGGATGAGAACCAAGTTTGTCTGCTGTCTACCAAGCAACCAATGCAAGGCA 1321
Qy 1203 TCTTCAATTTTACCCAGAGGCGGAGTGGCTGTGGGCTGTAGCGCTGACCTGGTCA 1262
Db 1322 TCTTTAACTGTATCCCAAGGAAAGGCGGATTCGCTGGGCTCGGATGCGGACGTGGTCA 1381
Qy 1263 TATGGAACCCCAAGGCGACCAAGATCATCTCTGCAAGACCCCAATCTGAAAGTGGAGT 1322
Db 1382 TCTGGGACCCCGCAAGTTGAGAACCAATTAACGCAAAAGTCAACAGTGGCGGCTGGAGT 1441
Qy 1323 ACAACATCTTTCGAGGAGTGGAGTGGCGGGAGCGCTGCGGTGCATTAAGTCAGGGCC 1382
Db 1442 ACAACATCTTTCGAGGATGAGAGTGCCACGCGCTCCCACTAGTGGTTCATCAGCAGGGCA 1501
Qy 1383 GAGTGGGCTGAGAGGAGCGGAGAGATGTTTGTGTACCCCGGGGGCGGCGCTTCGTCCCTC 1442
Db 1502 AGATCGCTTTTGAAGACGGAACATCAACGCTCAACAGGGGCATGGGCGGCTTCATTCGCG 1561
Qy 1443 GGAAGAACTTCCCGGACTTTGTCTACAAGAGGATCAAGCTCCGACACAGGCTGCGCGAGA 1502
Db 1562 GGAAGGGGTTTCCCGAGGACCTGTACCGGCGGTCAAAATCAGGAATTAAGGTTTGTGAT 1621
Qy 1503 TCACCGGTGTGCCCGGTGGGCTGTATGACGGGCGCGTCCACGAGGTGATGTGCTCCGCCA 1562
```

```
Db 1622 TGCAAGGGGTTTTCAGGGGCATGTATACGGTCTCTGTGTACGAGGTACAGGCTACACCCA 1681
Qy 1563 AGCCAGGAGTGGCGTCCGGCCCGGCGTCTGCCAGGCAAGATCTCGTGCCTCTCG 1622
Db 1682 AATATGCAACTCCCGCTCTTTCAGCCAAATCTTCGCTTTCTAAACACCAAGCCCCACCCA 1741
Qy 1623 TGGCAACCTACATCATGTCGGGTTTCAGGCTATCTGGGTCTCAGGCTGATGACCAATCG 1682
Db 1742 TCAGAAACCTCCACAGTCCAACTTCAGCTTATCAGTCCCCAGATAGATGACAACATC 1801
Qy 1683 CCGAGCGC 1690
Db 1802 CCAGGCGC 1809
```

## RESULT 8

```
US-09-949-016-4321
; Sequence 4321, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 4321
; LENGTH: 5032
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-4321
```

```
Query Match 50.3%; Score 850.6; DB 4; Length 5032;
Best Local Similarity 70.0%; Pred. No. 2.4e-188;
Matches 1160; Conservative 0; Mismatches 494; Indels 3; Gaps 1;
```

```
Qy 18 CCCCCAGGAGCAGGATGTCCTTCCAGGCGCAAGAAAGCATCCCCCGATCACAGGTGACC 77
Db 82 CCACCAACCCCACTGTCTTACCAAGGCAAGAAAGCAATCCCGCGATCAAGGTGACC 141
Qy 78 GCCTTCTGATCAGAGTGGGAGGATCGTGAATGACGACCAAGTCTTTTACGCTGATGTGC 137
Db 142 GTCTCTTATCAAGGAGGAGCAGAAATGTCATGATGATCAGTCTTTTATGCTGATTT 201
Qy 138 AGTGAAGATGGCTTTGATATAAAACAAATCGAGAAAACCTCATGCTCCCTGGGGGATCA 197
Db 202 ACATGAAGATGGCTTTATAAAACAAATTCGAGACAAATCTGATTTGTTCTTGGAGGATGA 261
Qy 198 AGACCAATTCAGCCCAAGCGCTGATGCTCTTCTGGTGGCGTTGACGTCCACACAAGGC 257
Db 262 AGACCAATTAAGGCAATGGGAAGATGGTATCTCTGGAGGATCGATGCTCATACTACT 321
Qy 258 TGCAAGTGCCTGTCTGGGCGATGACACCGGCTGACGACTTCTGTCAGGGGACCAAGGCGAG 317
Db 322 TCCAGATGCCATATAAGGGAATGACCAAGTAGATGACTTCTTCAAGGGACAAAGGGCG 381
Qy 318 CGCTAGCAGAGGAACCAACATGATCTTGGACCACTCTTTCCTCCGACACAGGGTGTGAGCC 377
Db 382 CCTTAGCAGTGGCACCAACATGATCATTTGACCATTTGCTGGTGGCTGAGCCTGAGTCCAGCC 441
Qy 378 TGCCTGGCGGCTTACGAGCAGTGGCGGAGGCGGCGGACAGCGGCGCTGCTGCGACTACT 437
Db 442 TGACTGAGGCTTATGAGAAATGGAGAGATGGGCTCATGGGAAGATTTGCTGTGACTATG 501
Qy 438 CCTGCACGTGGACATCAACCGATGGCATGAGAGCATCAAGGAGGAGCTGGAGGCCCTGG 497
```



Db 577 TCAAGGACAAAGGGTTAACTCTTATGTTTATATGCTTATAGGATTTGTATCAAG 636  
Qy 558 GCAGCAGCAGCCAGATGTACGAGATCTTACGACATATCCGGGACCTGGGGGCCCTTGGCCC 617  
Db 637 TATCTAACACAGAGCTCTATGAGATCTTACCTGCCTGGGAGAGCTGGGGGCCCATTTGCTC 696  
Qy 618 AGGTGACCGCTGAGAACGGGGACATCGTGGAGGAGGACAGAGCGGTTGCTGGAGCTCG 677  
Db 697 AAGTTTCATGCTGAGATGGGGATATCATTTGCCAGGAGCAAAACCCGCATGTTGGAAATGG 756  
Qy 678 GCATCACTGGCCCCGAGGGCCAGCTGCTCAGCCACCCCGAGGAGGTGGAGGCTTGAGCGG 737  
Db 757 GGNATACTGGCCAGAGGCCATGTACTGACGAGCCAGAGAGCTGGAGAGCTGAGGCTG 816  
Qy 738 TGTACCGAGCTGTCAACATCGCCAGCAGGCAAACTGCCCGCTGTACGTCAACCAAGGTGA 797  
Db 817 TGTTCGGTGCATCAGCATTTGCCAGCAAAACCAATTTGCCCTCTCTACGTCAACAAAGGTCA 876  
Qy 798 TGAGCAAGGGGGCGGCGGAGCGCATCGCTCAGGCCCAAGCCAGCGAGGGGTGGTGGTTTG 857  
Db 877 TGAGCAAGAGTGCAGCTGACCTCATCTCAAGCCAGGAAAAAGGAAATGTAGTCTTTG 936  
Qy 858 GGGAGCCCATCACCGCAGCTGGGCAACGACGGTTTCACTACTGGAGCAAGAACTGGG 917  
Db 937 GTGAGCCCATCACTGCAGCGCTCGGATAGTAGTGAACCCATTTATTGGAGCAAGAACTGGG 996  
Qy 918 CCAAGGCTGCAGCTTTCGTACATCAACCCCTGTCAACCCAGACCCCAACCAAGGAGCC 977  
Db 997 CCAAGGCGCTGCAATTTGTGACATCCCAACCCCTGAGCCCTGACCCAACTACTCCGACT 1056  
Qy 978 ACCTACCTGCTGCTGTCCAGGGGGAGCTTCAGGTGACAGCGAGCGCCCACTGCACCT 1037  
Db 1057 ACATCAACTCTTGTGTGGCAGCGGGGATCTGCAGCTATCTGGGAGTGCCCACTGCACCT 1116  
Qy 1038 TCACCACTGCCAGAGAGGCTGTGGGCAAGGACAACTTTCGGCTGATCCCGAGGGCACCA 1097  
Db 1117 TCAGCACTGCCAGAAAGCAATTTGGGAGGACAACTTTCAGCCATTCCTGAGGGCACCA 1176  
Qy 1098 ACGGCATTGAGGAGCGGATGTGATGGTCTGGAGAAATGTGTGGCTCTTGGGAAAGTGG 1157  
Db 1177 ATGGTGTGGAGAGCGGATGTGTGTCATCTGGGACAAAGGTGTGTGGCCACAGGGAAAAATGG 1236  
Qy 1158 ACGAGATAGTTCGTGGGTGACAGTACAAATCTGTCGCAAAATCTTCAATTTTACC 1217  
Db 1237 ACGAAACCAAGTTCTGTGGTGTGACAAAGCAAAACGCTGCAAGATCTTCAACCTGTATC 1296  
Qy 1218 CAAGGAAGGGCGAGTGGCTGTGGGCTCTGACGCTGACCTGTGTCATATGGAACCCCAAGG 1277  
Db 1297 CCGGCAAGGGAGAAATATCTGTGGGTCTGACAGCGACCTGCTCATCTGGGATCCAGATG 1356  
Qy 1278 CCACCAAGATCATCTCTGCAAGACCCCAATCTGAAACGTGGAGTACAAATCTTCGAGG 1337  
Db 1357 CTGTGAAGATCGTCTCTGCCAAGAACCAACCAAGTCTGCGGACAGATGACAAATCTTTGAAG 1416  
Qy 1338 GAGTGAAGTCCGGGAGCGCTGCGGTGTGTAAGTCAAGGCCCGAGTGGCGCTGGAGG 1397  
Db 1417 GGTGAGAGCTGCGCGGGGCTCTCTGCTGTGTCTATCTGCCAGGGCAAGATCATCTGGAAG 1476  
Qy 1398 ACGGGAAGATGTTTGTACCCCGGGGGCGGGCGCTTCTGTCCTCGGAAAAACATTCGCCG 1457  
Db 1477 ATGGCAACCTGACAGTGAACCGAGGGGCTGGCGCTTTCATACCTGAGCGCGCTTCTCGG 1536  
Qy 1458 ACTTTGTCTACAAGAGGATCAAAAGCTCGCAACAGGCTGGCGGAGATCCACGTTGCCCCC 1517  
Db 1537 ACTATGTCTACAAGCGCATTAAGCAACGGAGGAAGATGGCAGACCTGCGATGCCGTCCCAA 1596  
Qy 1518 GTGGGCTGTATGACGGGCCCGTCCAGAGGTGATGTGCTGCTGCAAGCCAGGAGTGGCG 1577  
Db 1597 GGGGCGATGTACGATGGGCGCTGTGTTGACCTGTGACCCACCCCAAC---AGGTGGCACCC 1653  
Qy 1578 CTCGGGCCCGCGCTCTCTGCCAGGCAAGATCTCTCGTGCCTCTCTGTGCGCAACTTACATC 1637

Db 1654 CCGCAGGCTCTGCTCGGGGCTCTCTCTACTCTCGGCCGAACCCACCTGTGAGGAATCTTTCATC 1713  
Qy 1638 AGTCGGGGTTACGCTATCTCTGGGTCTCAGGCTGATGA 1674  
Db 1714 AGTCGGGATTTAGCCTGTTCAGGCACCCCAAGTGGATGA 1750  
RESULT 10  
US-09-976-594-131  
; Sequence 131, Application US/09976594  
; Patent No. 6673549  
; GENERAL INFORMATION:  
; APPLICANT: Furness, Michael  
; APPLICANT: Buchbinder, Jenny  
; TITLE OF INVENTION: GENES EXPRESSED IN C3A LIVER CELL CULTURES TREATED WITH STEROIDS  
; FILE REFERENCE: PA-0041 US  
; CURRENT APPLICATION NUMBER: US/09/976,594  
; CURRENT FILING DATE: 2001-10-12  
; PRIOR APPLICATION NUMBER: 60/240,409  
; PRIOR FILING DATE: 2000-10-12  
; NUMBER OF SEQ ID NOS: 1143  
; SOFTWARE: PERL Program  
; SEQ ID NO 131  
; LENGTH: 3331  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc feature  
; OTHER INFORMATION: Incyte ID No. 6673549 408751.3  
US-09-976-594-131

Query Match 30.9%; Score 521.8; DB 4; Length 3331;  
Best Local Similarity 59.0%; Pred. No. 8.1e-112;  
Matches 951; Conservative 0; Mismatches 657; Indels 5; Gaps 3;  
Qy 69 CQAGTGACCGCTCTTGATCAGAGGTGGAGGATCGTGAATGACGACCAAGTCTCTTTTACG 128  
Db 66 CCAGCGTGAGGATCTCATCAAGGGAGGCAAGGTGGTGAACGATGACTGCACCCACGAG 125  
Qy 129 CTGATGTGCACTGGGAAGTGGCTTGATAAAACAAATCGGAGAAACCTCATGTCCTCTG 188  
Db 126 CTGACGTCTACATCAGAAATGGCATCATCCAGCAGGTGGCGCGAGCTCATGATCCCTG 185  
Qy 189 GGGGCATCAAGACCATTTGACGCCACCGCTCATGCTCTCTCTGTCGGTGGTTCAGCTCC 248  
Db 186 GCGGGGCCAAGGTGATGATGCCACAGGAAACTGGTGTATCCCTGGTGGATCGACACCA 245  
Qy 249 ACACAAAGGCTGCAGATGCTCTCTGGGCATGACACCGGCTGACGACTTCTGTCAAGGCA 308  
Db 246 GCACCCACTTCCACAGACTTTCATGAATGCCACGTCGTGGACGACTTCTACCATGGGA 305  
Qy 309 CCAAGCAGCGCTAGCAGGAGAAACCAATGATCTTGGACCAACGCTTCTCCCGCACCGG 368  
Db 306 CCAAGCAGCACTCTGTCGGAGGACCAACCATGATCATCGGCCACGCTCTGCGCCGACAAG 365  
Qy 369 GTGTGAGCTGTGCGGCGCTACGAGCAGTGGCGGAGCGGGCGGAGCAGCGGGCTGTCT 428  
Db 366 AGACCTCTCTTGGACGCTTATGAGAAAGTGGCGAGGTCTGGCCGACCCCAAGGTCTGCT 425  
Qy 429 GCGACTACTCTGACAGTGGACATCACCCTGATGGCATGAGAGCATCAAGAGGAGCTGG 488  
Db 426 GTGATTACGCCCTCCACGTGGGATCACCCTGTGGGACCCCAAGGTGAAGCAAGTGG 485  
Qy 489 AGGCCCTGTCAAGGAGAAAGGTGTGAATCTCTCTCTGTTCTTTCATGGCATCAAGGACC 548  
Db 486 AGACACTGGTGGAGGAAGGGTGTCAACTCTGTTCCAGATGTTTCATGACCTACAAAGACC 545  
Qy 549 GGTGCCAGTGCAGCAGCCAGCATGTAAGATCTTTCAGCATCTCCGAGCATCTCCGAGCTGGGG 608  
Db 546 TGTATCATGCTTCGAGACAGTGAAGCTGTACCAAGTGTTCACGCTTGCACGCTTCAAGGACATTGGGG 605  
Qy 609 CTTTGGCCCGCAGTGCACGCTGAGAACCGGAGCATCGTGGAGGAGGACAGAGCGGTTGC 668

Db 606 CAATCGCCCGGCTCCATGCTGTGAAATGGGGAGCTTGTGGCGGAGGGTGCTAAGAGGCGAC 665  
Qy TGGAGCTCGGCATCACTGGGCCCCGAGGGCCACGCTGCTCAGCCACCCCGAGGAGGTGAGG 728  
Db TGGATTTGGGGATCACAGGCCAGAGGAATCGAGATCAGCCGCTCCAGAGGAGCTGGAAG 725  
Qy CTGAGGGCGGTATCCGAGCTGTCTACATCGCCAAAGCAGGCAAACTGCCCCGCTGTAGCTCA 788  
Db CTGAAGCCACTCATCTGCTGTATACCAATGCAACAGGACTCACTGTCTCAATCTACCTGG 785  
Qy CCAAGGTGATGAGCAAGGGGGCGCCGACGCCATCGCTCAGGCCCAAGCGCAGAGGGGTGG 848  
Db TCAAGCTGTCACTATCTCGCTGTGAGCTGTATCGCAGCTGTAAAGATGCAAGAGGAGG 845  
Qy TCGTGTGTGGGAGCCCATCAACGCCAGCCCTGGGCAACCGAGCTTCACTACTCTGAGCA 908  
Db TTGTGCTGGCGGAGACCACTGACATGCCATGCCGCTGACAGGCTTACACTACTACCAAC 905  
Qy AGAATGGGCCAAGGCTGAGGCTTGTGCTAC-ATCACCCCTGTCAACCCAGAGCCCAAC 967  
Db AGAATGGTCCAGCGGCTGCTTATGTCAAGGTGCTCGCTCGCTGAGACTGGACACCAAC 965  
Qy ACGCAGACCACTCACTGCTGTCTGCTCCAGCGGGGACCTCCAGGTGACAGGCGAGCGCC 1027  
Db ACCTCAACCTCACTGAGCTGCTGCTGCTCAATGACACTCTGAACATCTGTGGCATCAT 1025  
Qy CACTGCACTTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1087  
Db CACGGCTTTCACCAAAATCAGAAAGCTATGGGCAAGGAGCTTTCACCAAGATCCCA 1085  
Qy GAGGCAACCAACGCTATGAGGAGCGCATGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1147  
Db CATGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1145  
Qy GGAAGATGATGAGAAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1207  
Db GGAAGATGATGAGAAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1205  
Qy AATTTTTCACCAAGAGGCGGAGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1267  
Db AACTGTATCCCGCAGGCGCGCATTTATCCCGGAGCGATGCTGATGTGTGTGTGTGTGT 1265  
Qy AACCCAGGCGCAACAGATCATCTCTGCAAGACCCACCAATCTGAACGTGGAGTACAAC 1327  
Db GAGCCAGAGGCCAACAGACCATCTCAGCCAGCAGCAGGCTCCAGGGAGGAGACTTCAAC 1325  
Qy ATCTTCAGGAGGTGAGTGGGAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1387  
Db CTGTATGAGAACATGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1385  
Qy GCGTGGAGCAGGGAAGATGTTGTCAACCGGGGGCGGCGCTTCTGCTCCCTCGGAA 1447  
Db GTGAT-GAGAACGGGCTTCTCATGTGCGCGGAGGGGCAACGGGCAAGTCTGCTCCCTGAGG 1444  
Qy ACATTCGGGACTTGTCTAAGAGGATCAAGAGCTCGCAACAGGCTGGCGGAGATCCAC 1507  
Db TCCTTCCAGACACTGTCTACAGAAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1504  
Qy GGTGTGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1567  
Db GGAAGTGAACCGCACTCCCTACCTGGGGGATGTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1564  
Qy GCGAGTGGCGCTCGGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1624  
Db GAGTGGGAACCCCACTGCGCAGACATCTTCAACCGGCGCGCTCACCGGCAATGGGGCATG 1624  
Qy CGCAACTCATAGTGGGGTTGAGCTTATGAGCTTATGAGCTTATGAGCTTATGAGCTTATG 1677  
Db AGGACCTTCAAGAAATCCAGCTTCAAGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1677

RESULT 11

US-09-252-991A-6970

; Sequence 6970, Application US/09252991A

; Patent No. 6551795

; GENERAL INFORMATION:

; APPLICANT: Marc J. Rubenfield et al.

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS

; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS

; FILE REFERENCE: 107196.136

; CURRENT APPLICATION NUMBER: US/09/252,991A

; CURRENT FILING DATE: 1999-02-18

; PRIOR APPLICATION NUMBER: US 60/074,788

; PRIOR FILING DATE: 1998-02-18

; PRIOR APPLICATION NUMBER: US 60/094,190

; PRIOR FILING DATE: 1998-07-27

; NUMBER OF SEQ ID NOS: 33142

; SEQ ID NO 6970

; LENGTH: 1524

; TYPE: DNA

; ORGANISM: Pseudomonas aeruginosa

; US-09-252-991A-6970

Query Match 23.0%; Score 389.2; DB 4; Length 1524;

Best Local Similarity 56.2%; Pred. No. 4.8e-81;

Matches 774; Conservative 0; Mismatches 598; Indels 6; Gaps 2;

Qy 79 CCTTCTGATCAGAGTGGGAGGATCGTGAATGACGACCACTCTTTTACGCTGATGCA 138

Db 90 CTTGTTGATCCGTGGGCGCCACCGTGTGTTACTTCAAGAGAGAGTTATCGAGCCGACGTCT 149

Qy 139 CGTGGAGATGCTTCAATAAACAATCGGAGAAACCTCATCGTCCCTGGGGGCAATCAA 198

Db 150 CTGTGCCAACGGCTCTATCCAGGCCATCGCGGAGAACTTCGAAACCCCTCCGCTGCGA 209

Qy 199 GACATTGACGCCACGCGCTGATGCTCTTCTGCTGGGCTTGAAGCTCCACACAGGCT 258

Db 210 CGTCTCGACGGGGTGGCCAGTACTGATCGCGGGGCGATCGACCCCATACCCACAT 269

Qy 259 GCAGATGCTGCTCTGGGATGACACCGGCTGACGACTTCTGTCAAGGCAACCAAGGCGC 318

Db 270 GCAGTTGCCCTTCATGGGACGCGTGGCCAGCGAGGACTTCTTCAGGGGCAACCGCGCGG 329

Qy 319 GCTAGCAGGAGGAAACCAACCATGATCTTGACACAGCTTCTCCCGCACACAGGCTGAGCCT 378

Db 330 GCTGGCGGAGGAGCAACCTCGATCATCGATTCCTGCTCATCCCCCAACCCGCGACAGTCT 389

Qy 379 GCTGGCGGCTTACGACGAGTGGCGGAGCGGCGGACAGCGCGCTGCTGCGACTACTC 438

Db 390 GCTGGAGGCTTCCACACCTTGGCGCGCTGGGCGGAGAAATCCGCTGCG---GACTACGG 446

Qy 439 CTTGCACTGGAACATCACCCGATGGCATGAGAGCATCAAGGAGGAGCTGGAGGCGCTTGT 498

Db 447 CTTCCACGTCGCCATCACCTGTTGGAGCGACGAGGTGCGCCGGGAAATGGGCGAGCTGTT 506

Qy 499 CAAGGAGAGGGTGTGAACTTCTTCTGCTTTCATGGCATCAAGGACCGGTGCCAGTG 558

Db 507 GGCACAGCAGCGGGTGAACAGCTTCAAGCATTCATGGGCTTACAGAACGCCCATCATGGC 566

Qy 559 CAGCGACAGCAGATGTAGAGATCTTACAGCATCTCGGACCTCGGGGCTTGGGCCCA 618

Db 567 CGCGGAGATACCTGTTGGTGGCGAGCTTCAGCGCTGCTGGAGCTGGGCGGCTGCCGAC 626

Qy 619 GGTGCACTGGAACCGGGGACATCGTGGAGGAGGAGCAAGCGGTGTCTGGAGCTCGG 678

Db 627 GGTGCACTGGAACCGGGGAACTGCTTCCACCTGCGAGCAAGAACTCTTCGCCAGGG 686

Qy 679 CATCACTGGGCGGAGGCGCATCGTGTCTAGCCACCCCGGAGGAGGTGGAGGCTGAGCGGT 738

Db 687 CTTACCGGCGGAGGCGCATCTCTGCTGCGGTCCGCGCAAGTCCAGGCGGAGGCGCGC 746

Qy 739 GTACCGAGCTGTCAACCATCGCAAGCAGGCAAACTGCGCGCTGTAGCTCACCAGGTGAT 798

Db 747 CAGCCGCGCATCCGCATCCCGAGAGCTGGGTGACGCGCTGTACTGTTGATATTTTC 806

Qy 799 GAGCAAGGGGCGGCGCCGACCCATCGCTCAGGCGCAAGGCGGAGGGTGTGCTGTTTGG 858

```
Db 807 CAGCCGCGAGGCGCTGGAAGATCGCTATGCCCCGCGCAAGGCGCCAGCGCTCTACGG 866
Qy 859 GGAGCCCATCAGCGCAGCTGTGGGCAACGAGGTTTCACTACTGAGGCAAGAACTGGGC 918
Db 867 CGAGTGTCTGGCGCGGCACTTGTGCTCGACGACAGCGTCTACCGGCAACCGGACTGGGC 926
Qy 919 CAGGCGTGAGCCTTGTGTCATACCCCTGTCAACCCAGACCCACACCGCAGACCA 978
Db 927 CACCGCGCGCGGATAGTATGAGCCCGCGTTTCCGTCCGTCGAACACAGGAAGCGCT 986
Qy 979 CTTCACTGTCTGTCTCAGCGGGGACCTTCAGAGTGACAGGCGAGCGGCCACTTCGACCTT 1038
Db 987 GTGGCGGCGCTCACTCCGGCA--ACCTGCATACACCGCACCGCACTGTCTT 1043
Qy 1039 CACCACTGCCAGAGGCTGTGGGCAAGACAACTTCGCGCTGATCCCGAGGGGCAACCA 1098
Db 1044 CTGCGCGAGCAGAGGCGCATGGGCGCGACGACTTCAGCAAGATTCCCAATGGCACGGC 1103
Qy 1099 CGGCATTGAGGAGCGCATGCGATGCTCTGGGAGAAATGTGTGGCTCTGGGAGATGGA 1158
Db 1104 CGGCATCGAGGACCGATGGCGCTGTCTGTGGGACGCGGGGTCAACAGCGGGCGCTGTG 1163
Qy 1159 CGAAGATGAGTTCTGTCGCGTGACCAAGTACAAATGCTGCCAAATCTTCAATTTTACCC 1218
Db 1164 GATGCAAGTGTCTGTCGCGTGACTCCACCAACACCGCAAGATCTTCACTGTTCCT 1223
Qy 1219 AAGGAAGGCGGAGTGGCTGTGGGCTGTGACGTGACCTGTGATATGAAACCCCAAGGC 1278
Db 1224 GCGCAAGGCGCGATCGCGTCCGCGCGGATGCGGACCTGTGCTCTGGGACCGCAGGG 1283
Qy 1279 CACCAAGATCATCTCTGCCAAGACCCACCAATCTGAACGTGGATCAACATCTTCGAGG 1338
Db 1284 CAGCCGACCTCTCGCGCGCCACCATCACAGCGGGTCAATTCAAATCTTCGAAGG 1343
Qy 1339 AGTGAAGTCCGGGAGCGCTCCGCTGTGTCATAAGTCAAGGCGGAGTGGCGCTGGAGGA 1398
Db 1344 TCGCACCGTGGCGGATNTCCAGCACCAACCATCAGCCAGGCAAGCTGTCTGGGCGC 1403
Qy 1399 CGGGAAGATGTTTGTACCCCCGGGGCGGCGCTTGTCTCCCTCGGAAACATTCGCG 1456
Db 1404 AGGCGACCTGGCGCGCAACCGCGCGGACGCTACGTGGAGCGCGCGCTTACCGG 1461
```

## RESULT 12

US-09-252-991A-6924  
; Sequence 6924, Application US/09252991A  
; Patent No. 6551795  
; GENERAL INFORMATION:

; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS

; FILE REFERENCE: 107196.136

; CURRENT APPLICATION NUMBER: US/09/252,991A

; CURRENT FILING DATE: 1999-02-18

; PRIOR APPLICATION NUMBER: US 60/074,788

; PRIOR FILING DATE: 1998-02-18

; PRIOR APPLICATION NUMBER: US 60/094,190

; PRIOR FILING DATE: 1998-07-27

; NUMBER OF SEQ ID NOS: 33142

; SEQ ID NO 6924

; LENGTH: 1611

; TYPE: DNA

; ORGANISM: Pseudomonas aeruginosa

US-09-252-991A-6924

Query Match 23.0%; Score 389.2; DB 4; Length 1611;  
Best Local Similarity 56.2%; Pred. No. 4.9e-81;  
Matches 774; Conservative 0; Mismatches 598; Indels 6; Gaps 2;

Qy 79 CTTCTGATCAGAGTGGGAGGATCGTGAATGACGACCAAGTCTCTTTTACCTGATGCA 138  
Db 79 CTTCTGATCGTGGGCGCACCGTGGTGTACTCAAGAGAGATTATCGAGCGGACGTGCT 138

```
Qy 139 CGTGGAAGTGTCTGATATAAAACAATCGGAGAAACCTCATCTGTCCTTGGGGGCATCAA 198
Db 139 CTGTGCAACGCGCTCATCCAGGCCATCGCGAGAACTTCGAAACCCCTTCGCGTCCGA 198
Qy 199 GACCAATTGACGCCCAAGCGCTCATGTCTTCTGTTGGCGTTGACGTCCACACAAAGGCT 258
Db 199 CTTCTCGACGGGGTGGCCAGTACTGTATGCGGGGCGCATCGACCCCTCATACCAT 258
Qy 259 GAGATGCTGTCTGGGATGACACCGGCTGACACTTCTGTCTGAGGGCACCAAGGCGC 318
Db 259 GCAATTGCTCTTCTAGGGACCGTGGCCAGCGAGGACTTCTTTCAGCGGCACCGCGCGCG 318
Qy 319 GTTACGAGAGGAAACCACTGATCTTGGACCAAGCTTCTCCCGACACAGGTTGTAGCCT 378
Db 319 GTTGGCGGAGGAGCACCTCGATCATCGACTTCTGTCATCCCCAACCCCGCAGACGCTG 378
Qy 379 GTTGGCGGCTTACGAGCAGTGGCGGAGCGGCGGACAGCGGCGCTGTGCGACTACTC 438
Db 379 GTTGGAGGCTTCCACACCTTGGCGGCTGGGCGCAGAAATCCGCTGCG--GACTACGG 435
Qy 439 CTTGCACTGTGGAATCACCCGATGGAATGAGAGCATCAAGAGGAGCTGAGGCTCTGCT 498
Db 436 CTTTCACTGCGCATCACCTGTGTGAGCAGCAGAGTTCGCGCGGAAATGGGCGAGCTGT 495
Qy 499 CAGGAGAGGCTGTGAATCTTCTGCTCTTCTGCTTCTGTCATGGCATACAGGACGGTGCAG 558
Db 496 GGCACAGCAGCGGGTGAAACAGCTTCAAGCATTCTATGGCTTCAAGAAACGCGCATATGGC 555
Qy 559 CAGCGACAGCCAGATGTACAGATCTTTCAGCATCATCCGCGGACCTTGGGGCGCTTGGCCCA 618
Db 556 CCGCGACGATACCTTGTGGCCAGCTTTCAGCGCTGCTGAGCTTGGGCGCGGTGCGGAC 615
Qy 619 GTTGCACGTGGAACCGGGAATCGTGTGAGGAGGAGAGAGCGGTTCTGTGAGTCTGG 678
Db 616 GTTGCACGCGGAGAACCGGGAATCTGCTTCTTCCACTGACGAGGAACTCTTCCGCCAGGG 675
Qy 679 CATCATGTCGCCCGGAGGCGCACGTGCTCAGCCACCCCGAGGAGGTGAGGCTGAGGCGGT 738
Db 676 CTTTACCGGCGCGGAGCGCATCTTCTGTGCGTCCGCGCAAGTCAAGGCGGAGGCGCGC 735
Qy 739 GTACCGAGCTGTCAACCATCGCCAAAGCAGGCAAACTGCCCCGCTGTACTCACCAGGTGAT 798
Db 736 CAGCGCGCATTCGATTCGCGAGACGTGGGTAGCCGCTGTACTTGTGCTATATTTC 795
Qy 799 GAGCAAGGGGGCGGCGCAGCCATGCTCAGGCAAGCGCAGAGGCGGTGCTGTGTTGG 858
Db 796 CAGCGCGAGGCGCTGAGCAGATCGCTTATGCGCGCGCAAGGGCGCAGCGGTCTACCG 855
Qy 859 GGAGCCCATCAGCGCGAGCTGGGACCGAGCGTTTCACTACTGAGCAAGAACTGGGC 918
Db 856 CGAGTGTCTGGCGCGGCACTGCTGTCTGACGACAGCGTCTACCGGCAACCGGAGCTGGGC 915
Qy 919 CAAAGCTGAGGCTTCTGTCACATCACCCCTGTCAACCCAGACCCCAACCGCAGACCA 978
Db 916 CACCGCGCGGATAGTGTATGAGCCCGCGTTCGCTCCGTCGAACACAGGAAGCGCT 975
Qy 979 CTTTCACTGCTGTGCTCAGCGGGGACCTTCAAGGTGACAGGCGAGCGCCACTGTGACCTT 1038
Db 976 GTGGCGGCGCTGCACTCGGCA--ACCTGATACACCGCCACCGACCACTGTGCTT 1032
Qy 1039 CACCATGCGCCAGAGGCTGTGGGCAAGGACAACTTTCGCGCTGATCCCGGCGGACCA 1098
Db 1033 CTGCGCGGAGCAGAGGCCATGCGGCGCGACGACTTCAGAAAGATTCCCAATGGCAGGC 1092
Qy 1099 CGGCATTGAGGAGCGCATGTGATGCTTGGGAGAAATGTGTGGCTCTCTGGGAAGTGA 1158
Db 1093 CGGCATCGAGGACCGCATGCGCTGTGTGGGACCGCGGGTCAACAGCGGGCGCTGTC 1152
Qy 1159 CGAAGATGTTGCTGCTCGCGGTGACGATCAAAATGTGCGCAAAATCTTCAATTTTACCC 1218
Db 1153 GATGCAAGTGTGCTGCGCTGACTCCACCTCCCAACACCGCGGAAGATCTTCAACCTGTTCC 1212
```

QY 1219 AAGGAGGGGCGAGTGGCTGTGGGCTGTGACGCTGACCTGTGTCATATGGAACCCCAAGGC 1278  
Db |||||  
QY 1213 GCGCAAGGCGCGATTCGCGTGGCGGATGCGGATGCGGCTGTGCTGTGGACCGCGAGG 1272  
Db |||||  
QY 1279 CACCAAGATCATCTCTGCAAGACCCACAACTCTGAAGTGAAGTACAACTCTTCGAGG 1338  
Db |||||  
QY 1273 CAGCGCACCTCTCTGCGCGCCACCATCACCAGGGTGTATTTCAACATCTTCGAAG 1332  
Db |||||  
QY 1339 AGTGAAGTGGCGGAGAGCGCTGCGCTGTGTCATAGTCAAGGCGGAGTGGCGCTGAGGA 1398  
Db |||||  
QY 1333 TCGCACCTGGCGCGCATTCACGACCCACACCATCAGCCAGGCAAGTGTCTCTGGCGCG 1392  
Db |||||  
QY 1399 CGGGAAGATGTTGTACCCCGGGGGCGCGCTGTGTCCTCGGAAACATTCCTCG 1456  
Db |||||  
QY 1393 AGGCGACTGGCGCGCAACCGCGCGCGAGCTAGCTGAGAGCGCGCGCTACCG 1450  
Db |||||

RESULT 13  
US-09-252-991A-6769/c  
; Sequence 6769, Application US/09252991A  
; Patent No. 6551795  
; GENERAL INFORMATION:  
; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: 107196.136  
; CURRENT APPLICATION NUMBER: US/09/252.991A  
; PRIOR FILING DATE: 1999-02-18  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR FILING DATE: 1998-07-27  
; NUMBER OF SEQ ID NOS: 33142  
; SEQ ID NO 6769  
; LENGTH: 1782  
; TYPE: DNA  
; ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-6769

Query Match 23.0%; Score 389.2; DB 4; Length 1782;  
Best Local Similarity 56.2%; Pred. No. 5.1e-81;  
Matches 774; Conservative 0; Mismatches 598; Indels 6; Gaps 2;

QY 79 CCTCTGATCAGAGGTGGAGGATCGTAATGACGACAGTCTCTTTAAGCTGATGCA 138  
Db CCTGTTGATCCGTGGCGCCACCGTGTACTCAGAGAGAGTTATCGAGCGAGTGCT 1601  
QY 139 CGTGAAGATGGCTTGTATATAAACAATCGAGAAACCTCATGTCCTGGGGGATCAA 198  
Db CTGTGCCAACGGCTCTCATCCAGGCCATCGCGAGAACCTCGAAACCCCTCCGGCTGCGA 1541  
QY 199 GACCATGAGCGCCACCGCTGATGCTCTTCTGTGGGGTGTGAGCTGCCACCAAGGCT 258  
Db CGTCTCTGACGGGTGGCCAGTACCTGATGTCGGCGGCGCATCGACCCCATCCACAT 1481  
QY 259 GCAGATGCTGTCTGCGCATGACACCGGTGACGACTTCTGTGAGGCAACCAAGGCG 318  
Db GCAGTTGCCCTTCATGGGACGGTGGCGACGAGGACTTCTCAGCGGCAACCGCGCGG 1421  
QY 319 GCTAGCAGGAGAACCAACATGATCTTGGACAGTCTTCCCGACACCGGTTGAGCT 378  
Db GCTGGCGGAGGAGCAACCTCTGATCATCGACTTCTGTCATCCCAACCCCGGACAGTGT 1361  
QY 379 GCTGGCGGCTTACGAGCAGTGGCGGAGCGCGGCGAGCGCGCTGCTGCGACTACTC 438  
Db GCTGGAGGCTTCCACACCTGGCGGCTGCGGCGAGAAATCCGCTGCG---GACTACGG 1304  
QY 439 CTTGACGCTGATCATCACCAGTGGATGAGAGCATCAAGGAGGAGCTGAGGCGCTGT 498  
Db CTTCCAGCTGCCATCACCCTGTGTGAGCGAGAGTGGCGGGAATGGCGAGTGT 1244  
QY 499 CAAGGAGAGGGTGTGAATCTCTTCTGCTTCTCATGGGCATCAAGAGCGGTGCCAGTG 558  
Db |||||

## RESULT 14

US-09-949-016-5326

; Sequence 5326, Application US/09949016

; Patent No. 6812339

; GENERAL INFORMATION:

; APPLICANT: VENTER, J. Craig et al.

; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED

; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF

; FILE REFERENCE: CL001307

; CURRENT APPLICATION NUMBER: US/09/949,016



CURRENT FILING DATE: 2000-04-14
PRIOR APPLICATION NUMBER: 60/241,755
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/237,768
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/231,498
PRIOR FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 207012
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 5326
LENGTH: 1909
TYPE: DNA
ORGANISM: Human
US-09-949-016-5326

Query Match
Best Local Similarity 61.1%; Pred. No. 5.2e-81;
Matches 661; Conservative 0; Mismatches 398; Indels 23; Gaps 1;

202 CATTTAGCCCGCCGCTGATGCTCTCTGCTGGGCTTGAAGTCCACACAGGCTGCA 261
9 CCTCGACGCGCGCGAAGCTGCTCTGCTGGGAGGATCGACACACAGCAGATGCA 68
262 GATGCTGCTCTGGGATGACACCGGCTGACGACTTCTGTGAGGCGACCAAGCAGCGCT 321
69 GTTCCCTTCATGGGCTCGGGTCCATCGAGACTTCCACACGCGCACCAAG----- 120
322 AGCAGGAGAACACACATGATCTTGACACAGCTCTTCCCGACACAGGCTGAGCCTGCT 381
121 -----CATGATTATTGATTTTGGCAATTCCTTCAGAAAGTGGCTCCCTCAT 165
382 GCGGCGCTACGACGCTGGCGGAGCGGCGGACGAGCGGCTGTGCGACTACTCCCT 441
166 TGAGGCGCTTCGACACTGGCGAAGCTGGGCTGATCCAAAGTTTGTGCGACTACAGCCT 225
442 GCACGTGGACATCACCGCATGGCATGAGAGCATCAAGAGGAGCTTGAGGCCCTTGTCAC 501
226 TCATGTGGCAGTACGCTGGTGGAGTGACACAGGTTAAAGAGAAATGAATCTTGTGCA 285
502 GGAGAGGGGTGAACTCTCTCTGCTGCTTTCATGGGATCAAGGACCGGTCGACAGTGCAG 561
286 AGATAAGGTGTTAACTCTTCAAGATGTTTATGGGCTATAAAGATCTGTACATGCTGAC 345
562 CGACGCCAGATGTCAGAGATCTTACGATCATCCGCGACCTGGGGCTTTGGCCAGGT 621
346 AGACCTGGAGCTGTACGAGCCTTCTCTCGGTGAAGAAATTTGAGCAATTTGCCAGGT 405
622 GCACGCTGAGAACCGGGACATCGTGGAGGAGGAGCAGAGCGGTTGTGAGCTCGGCAT 681
406 CCATGGGAAATGGAGACTTAATTCGACAGGGGACCAAGATGTTGGCTCTGGGGAT 465
682 CATGGCCCGAGGGGCACTGCTGACGCAACCGGAGGAGGTGAGGCTGAGCGGTGTA 741
466 AACAGGCGCTGAGGGGCACTGCTGCGCGCCAGAGGAGTGGAGGAGGAGGCGCACGCT 525
742 CCAGCTGTCAACCATCGCAGGAGCAAACTGCGCTGTACGTACCAAGGTGATGAG 801
526 GAGAGCCATCACCATAGCCAGCGCTGGAATGTCCTCTTACATTTGCTGATGATGAG 585
802 CAAGGGGGCGCGGACCGCCATCGCTCAGGCGCAAGCGCAGAGGGGTGCTGTTTGGGA 861
586 CAAGCTGTCAGCTAAGGTGATAGCGATGCAAGGAGAGATGGGAAGTGTCTATGTTGA 645
862 GCCCATACCGCCAGCCTGGGCGACCGGCTTCACTACTGAGGAGAGAACTGGGCGAA 921
646 ACCCATAGCAGCCAGTCTTGGCAGAGATGGCACTCACTACTGGAATAAAGATGGCACCA 705
922 GGCTGACGCTTGTGTCATATCACCCCTGTCAACCCAGAGCCCAACGAGCAGACACCT 981
706 TGAGGCCCACTGTCATGGGTCACTTTGGACCGAGACCCCTCAACACCGACTTCT 765
982 CACCTGCTGTGTCAAGCGGGAGCCTTCCAGGTGACAGGAGCGGCCCACTGCACTTTCAC 1041

766 CATGAATCTGTTGGCTAATGATGATCTAACCAACAGGAGTGAATCACTGCACTTTCAA 825
1042 CACTGCCAGAGGCTGTGGCAAGCAAACTTCGGCTGATCCCGAGGCGACCAACGG 1101
826 CACCTGCCAGAAAGCTCTTTGGGAAGGATGATTTTACCAAGATCCCAATGGGTGAATGG 885
1102 CATTGAGGAGCGCATGTCGATGCTGTGGGAGAAATGTGTGGCTCTTGGGAAGATGGACGA 1161
886 TGTGAGAGATCGATGTCGTAATGGAAGAAAGCGTGCATAGTGTGTAATGATGA 945
1162 GAATGAGTTCTGCGGCGTGCACAGTACAAATGCTGCAAAATCTTCAATTTTACCAAG 1221
946 AAACAGATTTGTGGCAGTTTACCAGCACAAATGCGACCAAAATTTTAAATCTCTATCAA 1005
1222 GAAGGGCGAGTGGCTGTGGCTCTGACGCTGACCTGTGATATGGAACCCCAAGGCCAC 1281
1006 AAAGGAAGAAATGCTGTAGGATGATGCTGACATTTGTTATTTGGGACCAAGGCCAC 1065
1282 CA 1283
1066 AA 1067

RESULT 15
US-09-734-237B-61
; Sequence 61, Application US/09734237B
; Patent No. 6818752
; GENERAL INFORMATION:
; APPLICANT: Rozzell, J. David
; APPLICANT: Bui, Peter
; APPLICANT: Hua, Ling
; TITLE OF INVENTION: SYNTHETIC GENES FOR ENHANCED EXPRESSION
; FILE REFERENCE: B583:40608
; CURRENT APPLICATION NUMBER: US/09/734,237B
; CURRENT FILING DATE: 2000-12-08
; PRIOR APPLICATION NUMBER: 09/494,921
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 79
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 61
; LENGTH: 1488
; TYPE: DNA
; ORGANISM: Pseudomonas putida
; US-09-734-237B-61

Query Match 21.5%; Score 363.8; DB 4; Length 1488;
Best Local Similarity 55.5%; Pred. No. 3.9e-75;
Matches 765; Conservative 0; Mismatches 607; Indels 7; Gaps 3;

79 CCTTCTGATCAGAGGTGGGAGGATCGTGAATGACGACCAAGTCTCTTTTACGCTGATGTGCA 138
6 CCTGTTGATCCGTGGCGCCACCGTGTCTACCCAGAGAGAGTATCCCGCCGATGTCTCT 65
139 CQTGGAAGATGGCTTGATAAACAAATCGGAGAAAACTCATCGTCCCTGGGGGATCAA 198
66 GTGTGTCGATGGCTCTGATCGTGCCATCGGCGCCAAACCTCGAACCGCCGACTGTGA 125
199 GACCATTTAGCCCGCCAGCGCTGATGCTCTTCTGTTGGGTTGAGCTCCACACAGGCT 258
126 AATCTCTGACGCGGAGCGGCGAGTACCTGATGCCCGCGGATCGACCCGATACCCACAT 185
259 GCAGATGCTCTCTGCGGATGACACCGGCTGACGACTTCTGTGAGGGCAACCAAGGCGC 318
186 GCAAGTTGCCATTATGCGGACCGTGGCGAGGAGTATTTCTTACGCGGACCCGAGCGG 245
319 GCTAGCAGGAGGAACCAACATGATCTTGGACCAAGTCTTCCCGACACAGGCTGTGAGCCT 378
246 CTTGCGGCGGCGACCAAGCTGATCATCGACTTCTGTCATCCCAACCCGAGGAGTCAAT 305
379 GCTGGCGGCTTACGAGCAGTGGCGGAGCGGCGGAGCGGCGCTGCTGCGGACTACTC 438
306 GCTGGAGGCTTCCACACCTGCGCGGCTGGGCGCAGA---AGAGCGCCAGGACTACGG 362

Qy	439	CCTGCAGTGGACATCAACCCGATGGCATGAGAGCATCAAGAGAGAGCTGAGAGCCCTGGT	498
Db	363	CTTCCAGGTTGCCATCACCTGGTGGAGCGAACAGGTGGCTGAAGAAATGGGCGAACTGGT	422
Qy	499	CAAGGAGAGGGTGTGAATCTCTTCTGGTCTTCAATGGCATACAAAGGACCGGTGCCAGTG	558
Db	423	AGCCAGCATGGGTGAACAGCTTCAAGCACTTCATGGCTTACAAAGATGCAATCATGGC	482
Qy	559	CAGCGACAGCCAGATGTACGAGATCTTTCAAGCATCATCGGACCTGGGGGCTTTGGCCCA	618
Db	483	CGCCGACGACACCTGGTGGCCAGCTTCGAGCGCTGCCCTGCAACTGGGTGCCGTGCCAC	542
Qy	619	GGTGCACTGTAGAAACGGGGACATCGTGAGGAGGAGAGAGCGGTTGCTGGAGCTCGG	678
Db	543	CGTGCACTGCCGAGAACCGGCGAACTGGTGTACCACTTCAGAAAAAACTGCTTGCACAGG	602
Qy	679	CATCACTGGGCCCGAGGGCCAGCTGCTCAGCCACCCCGAGGAGGTGGAGGCTGAGGCGGT	738
Db	603	CATGACCGGACGAGAGGCTCACCCCTTTGCGCCCTTCAAGTGAAGGTGAAGCGGC	662
Qy	739	GTACCGAGCTGTCAACCATCGCAAGCAGGCAAACTGCCCGCTGTACCTCACCAGGTGAT	798
Db	663	CAGCGCGCCATCGTATTCGCGAAACCATTTGGTACGCGCTGTATGTGTGACATTTTC	722
Qy	799	GAGCAAGGGGGCGGCCCGACGCTATCGCTCAGGCCCAAGCGAGAGGGGTGTGCTGTTGG	858
Db	723	CAGCGGTGAAGCACTGATGAATCACTATGCAAGCGGCCAAGGGCCAGCGGTTTACGG	782
Qy	859	GGAGCCCATCAGCCGAGCTGGGCAAGAGCGGTTCACATCTGGAGCAAGAACTGGGC	918
Db	783	CGAAGTCTTGGCCCGCCACCTGTGCTGGAGCAGAGGCTCTACCGTACCCGAGACTGGGC	842
Qy	919	CAAGGCTGCAGCCTTGTGTACATCACCCCTGTCAACCCAGAGACCCACACGCGCAGACCA	978
Db	843	CATGCGCTGGCTACGTGATGAGCCGCGCTTCCGCGCGGAGCACCCAGGAGGCGCT	902
Qy	979	CCTCACTGTCTGTGTCAGCGGGGACCTCCAGGTGACAGGCGAGCGCCCACTGCACCTT	1038
Db	903	GTGGCGGGTTGCAGTTCGGGCA---ACCTGCACACACCGGCCACCGACCACTGCTGTTT	959
Qy	1039	CACCACCTGCCAGAGGCTGTGGGCAAGGACAATCTCGCGCTGATCCCGAGGGGCACAA	1098
Db	960	CTGGCGCAACAGAAAGCATGGGCCCGGAGCACTTCAGTCCGATCCCCCAACCGGACCGC	1019
Qy	1099	CGGATTTGAGGAGCGCATGTGATGTGTCGGGAGAAATGTGTGGCTCTCGGGAAGATGGA	1158
Db	1020	CGGCATCGAAGACCGCATGGGGGTGCTGTGGATGCCGGTGTCAAGCGGGCGCCTGTC	1079
Qy	1159	CGAAGATGATGTCGTGCGGTGACAGTACAAATGTGCGCAAAATCTTCAATTTTACCC	1218
Db	1080	GATGCATGATGTTGTTGCGCTGACCTCCACACACGCGCAAAATCTTCAACCTTTTCCC	1139
Qy	1219	AAGGAGGGGCGAGTGGCTGTGGCTGTGACGCTGACCTGCTCATATGGAACCCCAAGGC	1278
Db	1140	ACGCAAGGGGCGCATCCGCGTGGGTGCGGACGCGCGACCTGCTGTGGGACCCGCGAGG	1199
Qy	1279	CACCAAGATCATCTCTGCCAAGACCCACAATCTGMAAGTGGAGTACAAATCTTTCGAGGG	1338
Db	1200	CATCGCACTCTATCGGCCAGACCCACACAGCGGTGGAATTCATATCTTTGAAG	1259
Qy	1339	AGTGAAGTCCGGGAGCGCTGCGGTGTCTAAGTCAAGTCAAGGCGGAGTGGCGCTGGAGGA	1398
Db	1260	CCGCACGTGCGGGGTCCCGAGCCACACCATCAGCCAGGGCAAGGTGCTCTGGGCGGA	1319
Qy	1399	CGGGAAGATG-TTGTGTCACCCCGGGGCGGCGCTTCGTCCCTCGGAAAAATTCGCG	1456
Db	1320	TGGCGACCTGCTGCGCGAGCGCGGGCGGGGTATGTGGAACGCGCGCGGTATCCG	1378

Search completed: September 24, 2005, 14:19:29  
Job time : 1106 secs

**THIS PAGE BLANK (USPTO)**

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 24, 2005, 07:02:36 ; Search time 1081 Seconds  
(without alignments)  
10452.922 Million cell updates/sec

Title: US-09-367-496C-7

Perfect score: 1690

Sequence: 1 ggcgcgccaccagagacc.....atgaccacatgccgcgacgc 1690

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 7400704 segs, 3343079526 residues

Total number of hits satisfying chosen parameters: 14801408

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications NA:\*

- 1: /cgn2\_6/ptodata/2/pubpna/US07\_PUBCOMB.seq.\*
- 2: /cgn2\_6/ptodata/2/pubpna/PCT\_NEW\_PUB.seq.\*
- 3: /cgn2\_6/ptodata/2/pubpna/US06\_NEW\_PUB.seq.\*
- 4: /cgn2\_6/ptodata/2/pubpna/US06\_PUBCOMB.seq.\*
- 5: /cgn2\_6/ptodata/2/pubpna/US07\_NEW\_PUB.seq.\*
- 6: /cgn2\_6/ptodata/2/pubpna/PCTUS\_PUBCOMB.seq.\*
- 7: /cgn2\_6/ptodata/2/pubpna/US08\_NEW\_PUB.seq.\*
- 8: /cgn2\_6/ptodata/2/pubpna/US08\_PUBCOMB.seq.\*
- 9: /cgn2\_6/ptodata/2/pubpna/US09A\_PUBCOMB.seq.\*
- 10: /cgn2\_6/ptodata/2/pubpna/US09B\_PUBCOMB.seq.\*
- 11: /cgn2\_6/ptodata/2/pubpna/US09C\_PUBCOMB.seq.\*
- 12: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq.\*
- 13: /cgn2\_6/ptodata/2/pubpna/US10A\_PUBCOMB.seq.\*
- 14: /cgn2\_6/ptodata/2/pubpna/US10B\_PUBCOMB.seq.\*
- 15: /cgn2\_6/ptodata/2/pubpna/US10C\_PUBCOMB.seq.\*
- 16: /cgn2\_6/ptodata/2/pubpna/US10D\_PUBCOMB.seq.\*
- 17: /cgn2\_6/ptodata/2/pubpna/US10E\_PUBCOMB.seq.\*
- 18: /cgn2\_6/ptodata/2/pubpna/US10F\_PUBCOMB.seq.\*
- 19: /cgn2\_6/ptodata/2/pubpna/US10G\_PUBCOMB.seq.\*
- 20: /cgn2\_6/ptodata/2/pubpna/US10H\_PUBCOMB.seq.\*
- 21: /cgn2\_6/ptodata/2/pubpna/US10I\_PUBCOMB.seq.\*
- 22: /cgn2\_6/ptodata/2/pubpna/US10\_NEW\_PUB.seq.\*
- 23: /cgn2\_6/ptodata/2/pubpna/US11A\_PUBCOMB.seq.\*
- 24: /cgn2\_6/ptodata/2/pubpna/US11\_NEW\_PUB.seq.\*
- 25: /cgn2\_6/ptodata/2/pubpna/US60\_NEW\_PUB.seq.\*
- 26: /cgn2\_6/ptodata/2/pubpna/US60\_PUBCOMB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1690	100.0	1690	9	US-09-986-632-9
2	1688.4	99.9	2741	20	US-10-723-860-5930
3	1686.8	99.8	2699	17	US-10-133-937-30
4	1686.8	99.8	2699	17	US-10-159-563-30
5	1686.8	99.8	2699	20	US-10-723-860-1486
6	1318.4	78.0	2352	21	US-10-498-788-60
7	980.2	58.0	2947	22	US-10-948-947A-48

8	965.4	57.1	1829	9	US-09-986-632-3	Sequence 3, Appli
9	965.4	57.1	4435	17	US-10-062-674-1971	Sequence 1971, Ap
10	963.8	57.0	4459	14	US-10-175-523-140	Sequence 140, App
11	963.8	57.0	4459	15	US-10-006-911-11	Sequence 11, Appl
12	963.8	57.0	4459	17	US-10-133-937-8	Sequence 8, Appli
13	963.8	57.0	4459	17	US-10-159-563-8	Sequence 8, Appli
14	963.8	57.0	4459	18	US-10-205-331-53	Sequence 53, Appl
15	963.8	57.0	4459	19	US-10-788-792-1	Sequence 1, Appli
16	963.8	57.0	4459	20	US-10-851-921-5	Sequence 5, Appli
17	963.8	57.0	4459	21	US-10-956-157-617	Sequence 617, App
18	963.8	57.0	4459	24	US-11-099-266-140	Sequence 140, App
19	963.8	57.0	5421	9	US-09-954-456-2178	Sequence 2178, Ap
20	963.8	57.0	5421	11	US-09-968-007A-432	Sequence 432, App
21	963.8	57.0	5421	15	US-10-171-581-178	Sequence 178, App
22	963.8	57.0	5421	15	US-10-006-911-3	Sequence 3, Appli
23	963.8	57.0	5421	19	US-10-776-827-66	Sequence 66, Appl
24	963.8	57.0	5421	19	US-10-775-169-34	Sequence 34, Appl
25	963.8	57.0	5421	20	US-10-807-308-1	Sequence 1, Appli
26	963.8	57.0	5421	21	US-10-843-641A-5205	Sequence 5205, Ap
27	963.8	57.0	5421	21	US-10-843-641A-5902	Sequence 6902, Ap
28	876.8	51.9	2869	15	US-10-084-817-351	Sequence 351, App
29	876.8	51.9	2928	17	US-10-133-937-33	Sequence 33, Appl
30	876.8	51.9	2928	17	US-10-159-563-33	Sequence 33, Appl
31	875.2	51.8	2842	9	US-09-986-632-7	Sequence 7, Appli
32	875.2	51.8	2842	9	US-09-736-457-1800	Sequence 1800, Ap
33	875.2	51.8	2842	9	US-09-902-941-1800	Sequence 1800, Ap
34	875.2	51.8	2842	9	US-09-849-626-1800	Sequence 1800, Ap
35	875.2	51.8	2842	14	US-10-017-754-1800	Sequence 1800, Ap
36	875.2	51.8	2842	16	US-10-113-872-1800	Sequence 1800, Ap
37	875.2	51.8	2842	17	US-10-283-017-1800	Sequence 1800, Ap
38	875.2	51.8	2842	21	US-10-852-335A-15	Sequence 15, Appl
39	873.4	51.7	1719	14	US-10-180-198-1	Sequence 1, Appli
40	850.6	50.3	5046	10	US-09-873-367C-102	Sequence 102, App
41	850.6	50.3	5046	21	US-10-843-641A-102	Sequence 102, App
42	850.6	50.3	5047	9	US-09-954-456-495	Sequence 495, App
43	850.6	50.3	5047	10	US-09-930-213-273	Sequence 273, App
44	850.6	50.3	5047	14	US-10-175-523-175	Sequence 175, App
45	850.6	50.3	5047	16	US-10-269-909-15	Sequence 15, Appl

#### ALIGNMENTS

##### RESULT 1

US-09-986-632-9 Application US/09986632  
; Sequence 9, Application US/09986632  
; Patent No. US2002011994A1  
; GENERAL INFORMATION:  
; APPLICANT: AGUERA, Michelle  
; TITLE OF INVENTION: Modulation of Ulip/CRMP activity for the prevention or  
; CURRENT APPLICATION NUMBER: US/09/986,632  
; FILE REFERENCE: P06974US01/BAS  
; PRIOR APPLICATION NUMBER: 2001-11-09  
; PRIOR FILING DATE: 2001-11-09  
; NUMBER OF SEQ ID NOS: 30  
; SOFTWARE: Patent In Ver. 2.1  
; SEQ ID NO 9  
; LENGTH: 1690  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-986-632-9

Query Match	100.0%	Score	1690	DB	9	Length	1690
Best Local Similarity	100.0%	Pred. No.	0				
Matches	1690	Conservative	0	Mismatches	0	Indels	0
QY	1	GGCGCCCTTACCAGAGACCCCGAGGAGGAGTGCTTCCAGGGCAAGAAAGCATCCC	60				
Db	1	GGCGCCCTTACCAGAGACCCCGAGGAGGAGTGCTTCCAGGGCAAGAAAGCATCCC	60				
QY	61	CGGATCAGGAGTGACCGCTTCTGTATCAGAGGTGGAGGATCGTGAATGACGACGATC	120				

[illegible]

1141	Db	GGCCTCTGGGAGATGGACGAGAAATGAGTTCTGTCGGGTGACCAAGTACAAATGCTGC	1200
1201	Qy	AATCTTCAATTTTTTACCCCAAGGAAGGGGAGTGGCTGTGGGCTCTGACGCTGACCTGGT	1260
1201	Db	AATCTTCAATTTTTTACCCCAAGGAAGGGGAGTGGCTGTGGGCTCTGACGCTGACCTGGT	1260
1261	Qy	CATATGGAACCCCAAGGCCCAAGAGATCATCTCTGCCAAGACCACAATCTGAACGTGGA	1320
1261	Db	CATATGGAACCCCAAGGCCCAAGAGATCATCTCTGCCAAGACCACAATCTGAACGTGGA	1320
1321	Qy	GTAACAACATCTTCGAGGGAGTGAAGTGC CGGGAGCGCTGCCGTGGTCTAATAGTCAAGG	1380
1321	Db	GTAACAACATCTTCGAGGGAGTGAAGTGC CGGGAGCGCTGCCGTGGTCTAATAGTCAAGG	1380
1381	Qy	CCGAGTGGCGCTGAGGACGGGAAGATGTTTGTGACCCCGGGGGCGGGCGCTTCGTCCC	1440
1381	Db	CCGAGTGGCGCTGAGGACGGGAAGATGTTTGTGACCCCGGGGGCGGGCGCTTCGTCCC	1440
1441	Qy	TCGGAAAAACATCTCCCGACCTTGTCTACAAGAGAGTCAAGCTCGCAACAGGCTGGCGGA	1500
1441	Db	TCGGAAAAACATCTCCCGACCTTGTCTACAAGAGAGTCAAGCTCGCAACAGGCTGGCGGA	1500
1501	Qy	GATCCAAGGTGTCCCGGTGGCTGTATGACGGGCCGCTCCACGAGGTGATGTGCTGTC	1560
1501	Db	GATCCAAGGTGTCCCGGTGGCTGTATGACGGGCCGCTCCACGAGGTGATGTGCTGTC	1560
1561	Qy	CAAGCCAGGGAGTGGCGCTCCGGCCCGGGGTCTCTGCCAAGCAAGATCTCCGTGGCTCC	1620
1561	Db	CAAGCCAGGGAGTGGCGCTCCGGCCCGGGGTCTCTGCCAAGCAAGATCTCCGTGGCTCC	1620
1621	Qy	TGTCGCAACCTACATCAGTCTGGGGTTCAGCCTATCTTGGGTCTTCAGGCTGATGACACAT	1680
1621	Db	TGTCGCAACCTACATCAGTCTGGGGTTCAGCCTATCTTGGGTCTTCAGGCTGATGACACAT	1680
1681	Qy	CGCCCGACGC	1690
1681	Db	CGCCCGACGC	1690

## RESULT 2

```

US-10-723-860-5930
; Sequence 5930, Application US/10723860
; Publication No. US20040253606A1
; GENERAL INFORMATION:
; APPLICANT: Aziz, Natasha
; APPLICANT: Gineburg, Wendy M.
; APPLICANT: Glotnik, Albert
; TITLE OF INVENTION: Methods of Diagnosis of Soft Tissue Sarcoma, Compositions &
; FILE REFERENCE: 05882.0193.NPUS01
; CURRENT APPLICATION NUMBER: US/10/723,860
; CURRENT FILING DATE: 2003-11-26
; PRIOR APPLICATION NUMBER: 60/429,739
; PRIOR FILING DATE: 2002-11-26
; NUMBER OF SEQ ID NOS: 8393
; SOFTWARE: Patentin version 3.2
; SEQ ID NO 5930
; LENGTH: 2741
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-723-860-5930

```

	Query Match	99.9%	Score 1688.4	DB 20	Length 2741
	Best Local Similarity	99.9%	Pred. No. 0		
	Matches 1689	Conservative 0	Mismatches 1	Indels 0	Gaps 0
Qy	1	GCGCCCTTACGAGAGACCCCGAGGAGCGATGTCTTCCAGGGCAGAAAAGCATCCC	60		
Db	118	GCGCCCTTACGAGAGACCCCGAGGAGCGATGTCTTCCAGGGCAGAAAAGCATCCC	177		
Qy	61	CCGATCAGAGTGACCGCCTTCTGATCAGAGTGGGAGGATCGTGNATGACACCAAGTC	120		
Db	178	CCGATCAGAGTGACCGCCTTCTGATCAGAGTGGGAGGATCGTGAATGACGACCAAGTC	237		

```
Qy 121 CTTTAAAGCTGATGCACTGGGAAGATGGCTTGATATAAAACAAATCGAGAGAAACCTCAT 180
Db 238 CTTTAAAGCTGATGCACTGGGAAGATGGCTTGATATAAAACAAATCGAGAGAAACCTCAT 297
Qy 181 CGTCCCTGGGGGATCAAGACCAATGACCGCCACCGGCTGATGGTCTTCTGGTGGGCT 240
Db 298 CGTCCCTGGGGGATCAAGACCAATGACCGCCACCGGCTGATGGTCTTCTGGTGGGCT 357
Qy 241 TGAGCTCCACAAAGGCTGAGATGCTGCTCTGGGATGACACCGGCTGACGACTTCTG 300
Db 358 TGAGCTCCACAAAGGCTGAGATGCTGCTCTGGGATGACACCGGCTGACGACTTCTG 417
Qy 301 TCAGGGCAACAAAGGCTGAGATGCTGCTCTGGGATGACACCGGCTGACGACTTCTG 360
Db 418 TCAGGGCAACAAAGGCTGAGATGCTGCTCTGGGATGACACCGGCTGACGACTTCTG 477
Qy 361 CGACCGGCTGAGCTGCTGGGGCTTACGAGCAGTGGGGAGCGGGGAGCAGCGC 420
Db 478 CGACCGGCTGAGCTGCTGGGGCTTACGAGCAGTGGGGAGCGGGGAGCAGCGC 537
Qy 421 GGCCTGCTGGGACTTACTCCCTGACGCTGAGCATCAACCGGATGGCATGAGAGCATCAAGGA 480
Db 538 GGCCTGCTGGGACTTACTCCCTGACGCTGAGCATCAACCGGATGGCATGAGAGCATCAAGGA 597
Qy 481 GGAGCTGGAGGCCCTGCTCAAGGAGAAAGGCTGTGAACTCTCTGCTGCTTCTGATGGCATA 540
Db 598 GGAGCTGGAGGCCCTGCTCAAGGAGAAAGGCTGTGAACTCTCTGCTGCTTCTGATGGCATA 657
Qy 541 CAAGGACCGGCTGAGCTGAGCGGACAGCCAGATGTAAGAGATCTTCAAGCATCAACCGGGA 600
Db 658 CAAGGACCGGCTGAGCTGAGCGGACAGCCAGATGTAAGAGATCTTCAAGCATCAACCGGGA 717
Qy 601 CTTGGGGGCTTGGCCAGAGTGACGCTGAGAAAGGAGATCGTGGAGGAGGAGCAGAA 660
Db 718 CTTGGGGGCTTGGCCAGAGTGACGCTGAGAAAGGAGATCGTGGAGGAGGAGCAGAA 777
Qy 720 GCGGTTGCTGGAGCTGGGCTCACTGCGCCCGAGGCGCACGCTGCTCAAGCCACCCCGAGGA 720
Db 778 GCGGTTGCTGGAGCTGGGCTCACTGCGCCCGAGGCGCACGCTGCTCAAGCCACCCCGAGGA 837
Qy 721 GGTGGAGGCTGAGCGGCTGACCGAGCTGTGACCATCGCCAGCAGGCAAACTGCCCGCT 780
Db 838 GGTGGAGGCTGAGCGGCTGACCGAGCTGTGACCATCGCCAGCAGGCAAACTGCCCGCT 897
Qy 781 GTACGTCACCAAGGTGATGAGAGGCGGCGCGACGCTGCTCAGGCAAGCGCAG 840
Db 898 GTACGTCACCAAGGTGATGAGAGGCGGCGCGACGCTGCTCAGGCAAGCGCAG 957
Qy 841 AGGGTGGTCTGTGTTGGGAGGCCATCACCGCCAGGCTGGGCAACGAGGTTCACTA 900
Db 958 AGGGTGGTCTGTGTTGGGAGGCCATCACCGCCAGGCTGGGCAACGAGGTTCACTA 1017
Qy 901 CTGAGCAGAACTGGGGCCAAAGGCTGACGCTTGTGATCAATCAACCCCTGTCAACCCAGA 960
Db 1018 CTGAGCAGAACTGGGGCCAAAGGCTGACGCTTGTGATCAATCAACCCCTGTCAACCCAGA 1077
Qy 961 CCCACCAACGAGCAGCAGCCTTCACTGCTGCTGCTCCAGCGGGAGCTCCAGGTGACAGG 1020
Db 1078 CCCACCAACGAGCAGCAGCCTTCACTGCTGCTGCTCCAGCGGGAGCTCCAGGTGACAGG 1137
Qy 1021 CAGCGCCCTGACACCTTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1080
Db 1138 CAGCGCCCTGACACCTTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1197
Qy 1081 GATCCCCGAGGCAACCAAGGCTATGAGAGGCGCATGTCATGCTGCTGGAGAAATGTGT 1140
Db 1198 GATCCCCGAGGCAACCAAGGCTATGAGAGGCGCATGTCATGCTGCTGGAGAAATGTGT 1257
Qy 1141 GGCTCTGGGAGAGTGAACGAGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1200
Db 1258 GGCTCTGGGAGAGTGAACGAGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1317
```

```
Qy 1201 AATCTTCAATTTTAAACCAAGGAAGGGCCAGTGGCTGTGGGCTCTCAACGCTGACCTGGT 1260
Db 1318 AATCTTCAATTTTAAACCAAGGAAGGGCCAGTGGCTGTGGGCTCTCAACGCTGACCTGGT 1377
Qy 1261 CATATGGAACCCCAAGSCCAACCAAGATCATCTGCGCAAGCAACCAACATCTGAACGTGGA 1320
Db 1378 CATATGGAACCCCAAGSCCAACCAAGATCATCTGCGCAAGCAACCAACATCTGAACGTGGA 1437
Qy 1321 GTACAACATCTTCCAGGGAGTGGAGTGCCTGGGGAGCGCTTGCCTGTGTCATAGTCAAGG 1380
Db 1438 GTACAACATCTTCCAGGGAGTGGAGTGCCTGGGGAGCGCTTGCCTGTGTCATAGTCAAGG 1497
Qy 1381 CCAGTGGCGCTGAGAGACGGGAAGATGTTTGTTCACCCCGGGGCGGGCCGCTTCTGTCCTC 1440
Db 1498 CCAGTGGCGCTGAGAGACGGGAAGATGTTTGTTCACCCCGGGGCGGGCCGCTTCTGTCCTC 1557
Qy 1441 TCGGAAAAACATCCCGGACTTTCTCTACAGAGATCAAGCTCGCAACAGGCTGGCGGA 1500
Db 1558 TCGGAAAAACATCCCGGACTTTCTCTACAGAGATCAAGCTCGCAACAGGCTGGCGGA 1617
Qy 1501 GATCCACGCTGTGCTCCCGCTGTATGACGGGCGCTTATGACAGAGTGTATGCTGCTGC 1560
Db 1618 GATCCACGCTGTGCTCCCGCTGTATGACGGGCGCTTATGACAGAGTGTATGCTGCTGC 1677
Qy 1561 CAAGCAGGAGTGGCGCTTCCCGCGCGCTCTGCGCCAGGCAAGATCTCCGTCCTCC 1620
Db 1678 CAAGCAGGAGTGGCGCTTCCCGCGCGCTCTGCGCCAGGCAAGATCTCCGTCCTCC 1737
Qy 1621 TGTGCGCAACCTACATCAGTTCGGGTTCAAGCTTCTGAGCTTCTGAGCTGATGACCAT 1680
Db 1738 TGTGCGCAACCTACATCAGTTCGGGTTCAAGCTTCTGAGCTTCTGAGCTGATGACCAT 1797
Qy 1681 CGCCCGCAGCG 1690
Db 1798 CGCCCGCAGCG 1807
```

## RESULT 3

```
US-10-133-937-30
; Sequence 30, Application US/10133937
; Publication No. US20030207278A1
; GENERAL INFORMATION:
; APPLICANT: Khan, Javed
; APPLICANT: Ringner, Markus
; APPLICANT: Peterson, Carsten
; APPLICANT: Meltzer, Paul
; TITLE OF INVENTION: METHODS FOR ANALYZING HIGH DIMENSIONAL DATA FOR CLASSIFYING,
; TITLE OF INVENTION: DIAGNOSING, PROGNOSTICATING, AND/OR PREDICTING DISEASES AND
; FILE REFERENCE: 11613.56US01
; CURRENT APPLICATION NUMBER: US/10/133,937
; CURRENT FILING DATE: 2002-11-04
; NUMBER OF SEQ ID NOS: 99
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 30
; LENGTH: 2699
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-133-937-30
```

Query Match 99.8%; Score 1686.8; DB 17; Length 2699;

Best Local Similarity 99.9%; Pred. No. 0;

Matches 1688; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
Qy 1 GCGCGCCCTTACAGAGACCCCGAGGAGCAGGATGCTCTCCAGGCGCAAGAAAGCATCCC 60
Db 118 GCGCGCCCTTACAGAGACCCCGAGGAGCAGGATGCTCTCCAGGCGCAAGAAAGCATCCC 177
Qy 61 CCGGATCACAGGTGACCGCTTCTGATCAGAGTGGAGGATCGTGAATCAGCAGCAGTC 120
Db 178 CCGGATCACAGGTGACCGCTTCTGATCAGAGTGGAGGATCGTGAATCAGCAGCAGTC 237
Qy 121 CTTTACGCTGATGTGCACTGTGAAGATGGCTTTGATAAAACAAATCGGAGAAACCTCAT 180
```

Db 238 CTTTACGCTGATGTCAGTGGAGATGGCTTGATATAACAAATCGAGAAAACCTCAT 297  
QY 181 CQTCCCTGGGGGATCAAGACCAATGACGCGCTGATGGTCTTCTCGTGGCGT 240  
Db 298 CQTCCCTGGGGGATCAAGACCAATGACGCGCTGATGGTCTTCTCGTGGCGT 357  
QY 241 TGACGTCACACAAAGGTCAGATGCTGTCCTGGGCATGACACGGCTGACGATTCG 300  
Db 358 TGACGTCACACAAAGGTCAGATGCTGTCCTGGGCATGACACGGCTGACGATTCG 417  
QY 301 TCAGGGCACCAAGGACGCTAGCAGAGGAAACACCATGATCTTGACCAACGCTTCCC 360  
Db 418 TCAGGGCACCAAGGACGCTAGCAGAGGAAACACCATGATCTTGACCAACGCTTCCC 477  
QY 361 CGACACGGGTGAGCTGCTGCGGCTTACGAGCAGTGGCGGAGCGGGCGACAGCG 420  
Db 478 CGACACGGGTGAGCTGCTGCGGCTTACGAGCGTGGCGGAGCGGGCGACAGCG 537  
QY 421 GGCCTGCTGGACTACTCCCTGACGCTGACATCACCGATGGCATGAGGATCAAGGA 480  
Db 538 GGCCTGCTGGACTACTCCCTGACGCTGACATCACCGATGGCATGAGGATCAAGGA 597  
QY 481 GGAGCTGGAGGCCCTGGTCAAGAGAAAGGTGTGAATCTCTTCTGCTTTCATGCGATA 540  
Db 598 GGAGCTGGAGGCCCTGGTCAAGAGAAAGGTGTGAATCTCTTCTGCTTTCATGCGATA 657  
QY 541 CAAAGCACCGGTGCAATGACGACGACAGCCAGATGTACGAGATCTTCAGCATATCCGGGA 600  
Db 658 CAAAGCACCGGTGCAATGACGACGACAGCCAGATGTACGAGATCTTCAGCATATCCGGGA 717  
QY 601 CTTGGGGGCTTGGCCAGGTGACACCTGAGAACGGGGACATCGTGGAGGAGGACGAA 660  
Db 718 CTTGGGGGCTTGGCCAGGTGACACCTGAGAACGGGGACATCGTGGAGGAGGACGAA 777  
QY 661 GCGGTTGCTGGAGCTCGGATCACTGGCCCGAGGCGCACGTGCTCAGCCACCCCGAGGA 720  
Db 778 GCGGTTGCTGGAGCTCGGATCACTGGCCCGAGGCGCACGTGCTCAGCCACCCCGAGGA 837  
QY 721 GGTGGAGGCTGAGCGGTGTACCGAGCTGTACCATGTGCAAGCAGGCAAACTGCCGCT 780  
Db 838 GGTGGAGGCTGAGCGGTGTACCGAGCTGTACCATGTGCAAGCAGGCAAACTGCCGCT 897  
QY 781 GTACGTCACCAAGGTGATGAGCAAGGGGGCGGCGGACATCGCTCAGGCGAAGCGGAG 840  
Db 898 GTACGTCACCAAGGTGATGAGCAAGGGGGCGGCGGACATCGCTCAGGCGAAGCGGAG 957  
QY 841 AGGGGTGGTGTGTTGGGAGGCCATCACCGCAGCCTGGGCACCGACGGTTTCACTA 900  
Db 958 AGGGGTGGTGTGTTGGGAGGCCATCACCGCAGCCTGGGCACCGACGGTTTCACTA 1017  
QY 901 CTGGAGCAAGAACTGGGCAAGGCTGACGCTTTCGTACATCACCCCTGTTCACCCAGA 960  
Db 1018 CTGGAGCAAGAACTGGGCAAGGCTGACGCTTTCGTACATCACCCCTGTTCACCCAGA 1077  
QY 961 CCCACCAAGGACGACCACTTCACTGCTGCTGCTCAGCGGGGACCTCCAGGTGACGG 1020  
Db 1078 CCCACCAAGGACGACCACTTCACTGCTGCTGCTCAGCGGGGACCTCCAGGTGACGG 1137  
QY 1021 CAGCGGCCACTGACCTTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1080  
Db 1138 CAGCGGCCACTGACCTTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1197  
QY 1081 GATCCCGAGGGGACCAACGCGCATTTGAGGAGCGCATGTGATGCTCTGGAGAAATGTGT 1140  
Db 1198 GATCCCGAGGGGACCAACGCGCATTTGAGGAGCGCATGTGATGCTCTGGAGAAATGTGT 1257  
QY 1141 GGCCTCTGGGAAGATGGAACGAGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1200  
Db 1258 GGCCTCTGGGAAGATGGAACGAGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1317  
QY 1201 AATCTTCAATTTTATCCAAAGGAAGGGCGAGTGGCTGTGGGCTCTGACGCTGACCTGCT 1260

Db 1318 AATCTTCAATTTTATCCAAAGGAAGGGCGAGTGGCTGTGGGCTCTGACGCTGACCTGCT 1377  
QY 1261 CATATGGAACCCCAAGGCCACCAAGATCATCTCTGCAAGACCCCAAACTCTGAACGTGA 1320  
Db 1378 CATATGGAACCCCAAGGCCACCAAGATCATCTCTGCAAGACCCCAAACTCTGAACGTGA 1437  
QY 1321 GTACAAACATCTTTCAGGGAGTGGAGTGCAGGGAGGCGCTGCGGTGCTATAAGTCAGGG 1380  
Db 1438 GTACAAACATCTTTCAGGGAGTGGAGTGCAGGGAGGCGCTGCGGTGCTATAAGTCAGGG 1497  
QY 1381 CCGAGTGGCGCTTGGAGGACGGGAAGATGTTGTCAACCCCGGGGCGGCTTGGTCCC 1440  
Db 1498 CCGAGTGGCGCTTGGAGGACGGGAAGATGTTGTCAACCCCGGGGCGGCTTGGTCCC 1557  
QY 1441 TCGGAAAACATTTCCCGGACTTGTCTACAGAGGATCAAAAGCTCGCAACAGGCTGGCGGA 1500  
Db 1558 TCGGAAAACATTTCCCGGACTTGTCTACAGAGGATCAAAAGCTCGCAACAGGCTGGCGGA 1617  
QY 1501 GATCCACGGTGTGCGCGTGTATGACGGGCGGCTCCACGAGGTGATGGTGCCTGC 1560  
Db 1618 GATCCACGGTGTGCGCGTGTATGACGGGCGGCTCCACGAGGTGATGGTGCCTGC 1677  
QY 1561 CAAAGCAGGAGTGGCGCTCGCGCGCGCTCTGCCCCAGGCAAGATCTCGTGCCTCC 1620  
Db 1678 CAAAGCAGGAGTGGCGCTCGCGCGCGCTCTGCCCCAGGCAAGATCTCGTGCCTCC 1737  
QY 1621 TGTGCGCAACCTACATCAGTCGGGTTTACGCTATCTGGGTCTCAGGCTGATGACCAT 1680  
Db 1738 TGTGCGCAACCTACATCAGTCGGGTTTACGCTATCTGGGTCTCAGGCTGATGACCAT 1797  
QY 1681 CGCCCGACGC 1690  
Db 1798 CGCCCGACGC 1807

## RESULT 4

US-10-159-563-30  
; Sequence 30, Application US/10159563  
; Publication No. US20040009154A1  
; GENERAL INFORMATION:  
; APPLICANT: Khan, Javed  
; APPLICANT: Ringner, Markus  
; APPLICANT: Peterson, Carsten  
; APPLICANT: Meltzer, Paul  
; TITLE OF INVENTION: SELECTIONS OF GENES AND METHODS OF USING THE SAME FOR  
; TITLE OF INVENTION: DIAGNOSIS AND FOR TARGETING THE THERAPY OF SELECT CANCERS  
; FILE REFERENCE: 11613.56US11  
; CURRENT APPLICATION NUMBER: US/10/159,563  
; PRIOR FILING DATE: 2002-12-09  
; PRIOR APPLICATION NUMBER: US 10/133,937  
; PRIOR FILING DATE: 2002-04-25  
; NUMBER OF SEQ ID NOS: 444  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 30  
; LENGTH: 2699  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-159-563-30

Query Match 99.8%; Score 1686.8; DB 17; Length 2699;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 1688; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GCGGCCCCCTACCAAGAGACCCCGAGGAGGATGTCTCTTCCAGGGGCAAGAAAGCATCCC 60  
Db 118 GCGGCCCCCTACCAAGAGACCCCGAGGAGGATGTCTCTTCCAGGGGCAAGAAAGCATCCC 177  
QY 61 CCGGATCAGAGTGAACCGCTTCTGATCAGAGTGGAGGATCGTGAATGACGACGATC 120  
Db 178 CCGGATCAGAGTGAACCGCTTCTGATCAGAGTGGAGGATCGTGAATGACGACGATC 237  
QY 121 CTTTACGCTGATGTGACGTGGAAGATGGCTTTGATATAAACAATCGGAGAAAACCTCAT 180



238 Db CTTTTCAGCTGATGTGCACTGGAGAGTGGCTTGATATAAACAATCGGAGAAAACCTCAT 297  
181 QY CGTCCCTGGGGGCAATCAAGACCAATGACCGCCACCGGCTGATGCTCTTCTGCTGGCGT 240  
298 Db CGTCCCTGGGGGCAATCAAGACCAATGACCGCCACCGGCTGATGCTCTTCTGCTGGCGT 357  
241 QY TGACGTCAACAAGGCTGAGATGCTCTGCTGGGCAATGACACCGGCTGACGACTTCTG 300  
358 Db TGACGTCAACAAGGCTGAGATGCTCTGCTGGGCAATGACACCGGCTGACGACTTCTG 417  
301 QY TCAGGGCAACNAGGCGCTAGCAGGAGGAACACCATGATCTTGGACCAAGCTCTTCCC 360  
418 Db TCAGGGCAACNAGGCGCTAGCAGGAGGAACACCATGATCTTGGACCAAGCTCTTCCC 477  
361 QY CGACACGGGTGTGAGCCTGTGCGGCTTACGAGCAGTGGCGGAGCGGGCGGACAGCGC 420  
478 Db CGACACGGGTGTGAGCCTGTGCGGCTTACGAGCAGTGGCGGAGCGGGCGGACAGCGC 537  
421 QY GGCTGTGCGACTATCTCCCTGCACTGTGGAATCATACCCGATGGCATGAGAGCATCAAGGA 480  
538 Db GGCTGTGCGACTATCTCCCTGCACTGTGGAATCATACCCGATGGCATGAGAGCATCAAGGA 597  
481 QY GGAAGTGGAGGCGCTGTGAGGAGAGGCTGTGAATCTCTGCTCTTCAATGGCAT 540  
598 Db GGAAGTGGAGGCGCTGTGAGGAGAGGCTGTGAATCTCTGCTCTTCAATGGCAT 657  
541 QY CAAGGACCGGTGCGAGTGCAGCGACAGCAGATGTACGAGATCTTCAAGCATCATCCGGGA 600  
658 Db CAAGGACCGGTGCGAGTGCAGCGACAGCAGATGTACGAGATCTTCAAGCATCATCCGGGA 717  
601 QY CTTGGGGGCTTGGCCAGGTGCACTGCTGAGAAACGGGGAATCTGTGGAGGAGGACAGAA 660  
718 Db CTTGGGGGCTTGGCCAGGTGCACTGCTGAGAAACGGGGAATCTGTGGAGGAGGACAGAA 777  
661 QY GCGGTGTGAGGCTGGGATCACTGCGCCCGAGGGGCACTGCTGAGCCACCCCGAGGA 720  
778 Db GCGGTGTGAGGCTGGGATCACTGCGCCCGAGGGGCACTGCTGAGCCACCCCGAGGA 837  
721 QY GGTGAGGCTGAGCGGTGTACCGAGCTGTACCATCGCCAAAGCAGCAAACTGCCCGCT 780  
838 Db GGTGAGGCTGAGCGGTGTACCGAGCTGTACCATCGCCAAAGCAGCAAACTGCCCGCT 897  
781 QY GTACGTCAACAGGTGATGAGCAAGGGGGCGGCGCAAGCCATCGCTCAGGCCAAGGCGCAG 840  
898 Db GTACGTCAACAGGTGATGAGCAAGGGGGCGGCGCAAGCCATCGCTCAGGCCAAGGCGCAG 957  
841 QY AGGGGTGTGCTTTGGGAGGCGCATCACCGGAGCTGGGCAACCGAGGTTCACTA 900  
958 Db AGGGGTGTGCTTTGGGAGGCGCATCACCGGAGCTGGGCAACCGAGGTTCACTA 1017  
901 QY CTGGAGCAAGAACTGGGCGCAAGGCTGCGAGCTTCTGTCACATCACCCCTGTCAACCCAGA 960  
1018 Db CTGGAGCAAGAACTGGGCGCAAGGCTGCGAGCTTCTGTCACATCACCCCTGTCAACCCAGA 1077  
961 QY CCCCACACGCGAGACCACTCACTGCTTGTGTCAGGCGGGAACCTCCAGGTGACAGG 1020  
1078 Db CCCCACACGCGAGACCACTCACTGCTTGTGTCAGGCGGGAACCTCCAGGTGACAGG 1137  
1021 QY CAGCGGCCATGTCACTTCACTGCGCCAGAGGCTGTGGGCAAGGCAAACTTGGCGCT 1080  
1138 Db CAGCGGCCATGTCACTTCACTGCGCCAGAGGCTGTGGGCAAGGCAAACTTGGCGCT 1197  
1081 QY GATCCCGAGGGGCAACCAAGGCTTGGAGGCGCATGTGCTGCTGGGAGAAATGTGT 1140  
1198 Db GATCCCGAGGGGCAACCAAGGCTTGGAGGCGCATGTGCTGCTGGGAGAAATGTGT 1257  
1141 QY GGCTCTTGGGAAGATGAGACGAGATGAGTTGCTGCGGCTGACCAAGTACAAATCTGCCAA 1200  
1258 Db GGCTCTTGGGAAGATGAGACGAGATGAGTTGCTGCGGCTGACCAAGTACAAATCTGCCAA 1317  
1201 QY AATCTTCAATTTTACCAAGGAGGGGCGAGTGGCTGTGGGCTCTGACGCTGACTGGT 1260  
1318 Db AATCTTCAATTTTACCAAGGAGGGGCGAGTGGCTGTGGGCTCTGACGCTGACTGGT 1377

1261 QY CATATGGAACCCCAAGGCCACCAAGATCATCTCTGCCAAGACCCCAATCTGAACGTGGA 1320  
1378 Db CATATGGAACCCCAAGGCCACCAAGATCATCTCTGCCAAGACCCCAATCTGAACGTGGA 1437  
1321 QY GTACAACATCTTCGAGGAGTGGAGTGCCTGGGAGCGCTTGCCTGGTTCATAAGTCAGGG 1380  
1438 Db GTACAACATCTTCGAGGAGTGGAGTGCCTGGGAGCGCTTGCCTGGTTCATAAGTCAGGG 1497  
1381 QY CCGAGTGGCGCTGAGGAGCGGGAAGATGTTTGTACCCCGGGGGCGGCGCTTCGTCCC 1440  
1498 Db CCGAGTGGCGCTGAGGAGCGGGAAGATGTTTGTACCCCGGGGGCGGCGCTTCGTCCC 1557  
1441 QY TCAGAAAAACAATCCCGGACTTTTGTCTACAAGAGGATCAAAAGCTCGCAACAGGCTGGCGGA 1500  
1558 Db TCAGAAAAACAATCCCGGACTTTTGTCTACAAGAGGATCAAAAGCTCGCAACAGGCTGGCGGA 1617  
1501 QY GATTCACGCTGTGCTGCGGCTGTATGACGGGCGCTTCAAGAGTGTATGCTGCTGC 1560  
1618 Db GATTCACGCTGTGCTGCGGCTGTATGACGGGCGCTTCAAGAGTGTATGCTGCTGC 1677  
1561 QY CAAGCCAGGAGTGGCGCTCCGCGCGCTCCGCGCGGCTCCGCGCGGCAAGATCTCCGTGCTCC 1620  
1678 Db CAAGCCAGGAGTGGCGCTCCGCGCGGCTCCGCGCGGCTCCGCGCGGCAAGATCTCCGTGCTCC 1737  
1621 QY TGTGCGCAACCTACATCAGTCCGGGTTTCAAGCTTCTGAGCTTCTGAGCTGATGACCAT 1680  
1738 Db TGTGCGCAACCTACATCAGTCCGGGTTTCAAGCTTCTGAGCTTCTGAGCTGATGACCAT 1797  
1681 QY CGCCCGACGC 1690  
1798 Db CGCCCGACGC 1807

## RESULT 5

US-10-723-860-1486  
; Sequence 1486, Application US/10723860  
; Publication No. US20040253606A1  
; GENERAL INFORMATION:  
; APPLICANT: Aziz, Natasha  
; APPLICANT: Ginsburg, Wendy M.  
; APPLICANT: Zlotnik, Albert  
; TITLE OF INVENTION: Methods of Diagnosis of Soft Tissue Sarcoma, Compositions &  
; TITLE OF INVENTION: Methods for Screening for Soft Tissue Sarcoma Modulators  
; FILE REFERENCE: 05882.0193.NPUS01  
; CURRENT APPLICATION NUMBER: US/10/723,860  
; PRIOR FILING DATE: 2003-11-26  
; PRIOR APPLICATION NUMBER: 60/429,739  
; NUMBER OF SEQ ID NOS: 8393  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 1486  
; LENGTH: 2699  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-723-860-1486

Query Match 99.8%; Score 1686.8; DB 20; Length 2699;

Best Local Similarity 99.9%; Pred. No. 0;

Matches 1688; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GCGCGCCCTACAGAGACCCCGAGGAGCAGGATGTCTTCCAGGGCAAGAAAAGCATCCC 60  
Db 118 GCGCGCCCTACAGAGACCCCGAGGAGCAGGATGTCTTCCAGGGCAAGAAAAGCATCCC 177  
QY 61 CCGGATCAGAGTACCGGCTTCTGATCAGAGTGGGAGGATCGTGAATGACGACCAAGTC 120  
Db 178 CCGGATCAGAGTACCGGCTTCTGATCAGAGTGGGAGGATCGTGAATGACGACCAAGTC 237  
QY 121 CTTTTCAGCTGATGTGCACTGGAAGATGGCTTGTGATAAAACAATCGGAGAAAACCTCAT 180  
Db 238 CTTTTCAGCTGATGTGCACTGGAAGATGGCTTGTGATAAAACAATCGGAGAAAACCTCAT 297

181 CGTCCCTGGGGGATCAAGACCAATTGACGCCCGCTGATGGTCTCTCCTGGTGGCGT 240  
Db |  
298 CGTCCCTGGGGGATCAAGACCAATTGACGCCCGCTGATGGTCTCTCCTGGTGGCGT 357  
Qy |  
241 TGAAGTCCACAAAGGCTGCGATGCTGTCTTGGGATGACACCGGCTGACGACTTCTG 300  
Db |  
358 TGAAGTCCACAAAGGCTGCGATGCTGTCTTGGGATGACACCGGCTGACGACTTCTG 417  
Qy |  
301 TCAGGACCCAGGAGCGCTAGCAGAGGAGAACCAACATGATCTTGGACCAAGCTCTTCC 360  
Db |  
418 TCAGGACCCAGGAGCGCTAGCAGAGGAGAACCAACATGATCTTGGACCAAGCTCTTCC 477  
Qy |  
361 CGACACGGGTGTGAGCGCTGTGCGCGCTACAGAGCAGTGGCGGAGCGGCGGAGCAGCGC 420  
Db |  
478 CGACACGGGTGTGAGCGCTGTGCGCGCTACAGAGCAGTGGCGGAGCGGCGGAGCAGCGC 537  
Qy |  
421 GGCCTGCTGGAATCTTCCCTGCACTGCTGCAATCAACCGGATGAGAGCATCAAGGA 480  
Db |  
538 GGCCTGCTGGAATCTTCCCTGCACTGCTGCAATCAACCGGATGAGAGCATCAAGGA 597  
Qy |  
481 GAGCTGGAGGCGCTGGTCAAGAGAGAGGTGTGAATCTCTTCTGGTCTTCAATGGCATA 540  
Db |  
598 GAGCTGGAGGCGCTGGTCAAGAGAGAGGTGTGAATCTCTTCTGGTCTTCAATGGCATA 657  
Qy |  
541 CAAGGACCGGTGCAAGTGCAGGACAGCAGCAGATGTACAGATCTTCAAGATCAATCCGGGA 600  
Db |  
658 CAAGGACCGGTGCAAGTGCAGGACAGCAGCAGATGTACAGATCTTCAAGATCAATCCGGGA 717  
Qy |  
601 CTTGGGGGCTTGGCCCAAGTGCAGTGCAGTGCAGTGCAGTGCAGTGCAGTGCAGTGCAGT 660  
Db |  
718 CTTGGGGGCTTGGCCCAAGTGCAGTGCAGTGCAGTGCAGTGCAGTGCAGTGCAGTGCAGT 777  
Qy |  
661 GCGGTTGCTGGAGCTGGGATCACTGGCCCCGAGGCGCAGTGTCTCAGCCACCCGAGGA 720  
Db |  
778 GCGGTTGCTGGAGCTGGGATCACTGGCCCCGAGGCGCAGTGTCTCAGCCACCCGAGGA 837  
Qy |  
721 GGTGGAGGCTGAGCGGTGTACCGAGCTGTACCATCGCCCAAGCAGCAAACTGCCCGCT 780  
Db |  
838 GGTGGAGGCTGAGCGGTGTACCGAGCTGTACCATCGCCCAAGCAGCAAACTGCCCGCT 897  
Qy |  
781 GTAAGTCAACAAAGTGTAGCAGAGGCGGCGCGCAGCGCATCGCTCAGGCGCAAGCGCAG 840  
Db |  
898 GTAAGTCAACAAAGTGTAGCAGAGGCGGCGCGCAGCGCATCGCTCAGGCGCAAGCGCAG 957  
Qy |  
841 AGGGGT 900  
Db |  
958 AGGGGT 1017  
Qy |  
901 CTGGAGCAAGAACTGGGCGCAAGGCTGTGAGCGCTTGTGTGTGTGTGTGTGTGTGTGTGTGT 960  
Db |  
1018 CTGGAGCAAGAACTGGGCGCAAGGCTGTGAGCGCTTGTGTGTGTGTGTGTGTGTGTGTGTGT 1077  
Qy |  
961 CCCCAACAGGCGAGACCACTTCACTTCACTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1020  
Db |  
1078 CCCCAACAGGCGAGACCACTTCACTTCACTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1137  
Qy |  
1021 CAGGCGCACTGCACTTCACTTCACTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1080  
Db |  
1138 CAGGCGCACTGCACTTCACTTCACTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1197  
Qy |  
1081 GATCCCGGAGGCGCAACAGGCTTGGAGGCGCATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1140  
Db |  
1198 GATCCCGGAGGCGCAACAGGCTTGGAGGCGCATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1257  
Qy |  
1141 GGCCTCTGGAGAGATGGAAGAGATGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1200  
Db |  
1258 GGCCTCTGGAGAGATGGAAGAGATGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1317  
Qy |  
1201 AATCTTCAATTTTTTACCAAGAGAGGCGGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1260  
Db |  
1318 AATCTTCAATTTTTTACCAAGAGAGGCGGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1377  
Qy |  
1261 CATATGGAACCCCAAGGCGCACCAAGATCATCTCTGTGCAAGACCCCAATCTGAAAGCTGGA 1320

1378 CATATGGAACCCCAAGGCGCACCAAGATCATCTCTGCAAGACCCCAATCTGAAAGCTGGA 1437  
Qy |  
1321 GTACAAATCTTTGAGAGGAGTGGAGTGCAGGAGAGGCGCTGCGTGGTGTATAGTCAAGG 1380  
Db |  
1438 GTACAAATCTTTGAGAGGAGTGGAGTGCAGGAGAGGCGCTGCGTGGTGTATAGTCAAGG 1497  
Qy |  
1381 CGAGTGGCGCTGGAGGAGCGGAGATGTTTGTCAACCCGCGGCGGCGCTTGTGTGTGTGT 1440  
Db |  
1498 CGAGTGGCGCTGGAGGAGCGGAGATGTTTGTCAACCCGCGGCGGCGCTTGTGTGTGTGT 1557  
Qy |  
1441 TCGGAAAAATTTCCCGGAGCTTTTGTCTACAAAGAGGATCAAAGCTCGCAACAGGCTGGCGGA 1500  
Db |  
1558 TCGGAAAAATTTCCCGGAGCTTTTGTCTACAAAGAGGATCAAAGCTCGCAACAGGCTGGCGGA 1617  
Qy |  
1501 GATCCACGGTGTGCGCGCTGAGGCTGTATGACGGGCGCGTCCAGAGGTGATGGTGTGTGTGT 1560  
Db |  
1618 GATCCACGGTGTGCGCGCTGAGGCTGTATGACGGGCGCGTCCAGAGGTGATGGTGTGTGTGT 1677  
Qy |  
1561 CAAGCAGGAGTGTGGCGCTCCGCGCGCTGCGCGCGCTGCGCGCGCAAGATCTCGTGTGTGTGT 1620  
Db |  
1678 CAAGCAGGAGTGTGGCGCTCCGCGCGCTGCGCGCGCTGCGCGCGCAAGATCTCGTGTGTGTGT 1737  
Qy |  
1621 TGTGCGCAACCTACATCATCATCATCATCATCATCATCATCATCATCATCATCATCATCATCAT 1680  
Db |  
1738 TGTGCGCAACCTACATCATCATCATCATCATCATCATCATCATCATCATCATCATCATCATCAT 1797  
Qy |  
1681 CGCCCGACGC 1690  
Db |  
1798 CGCCCGACGC 1807

## RESULT 6

US-10-498-788-60  
; Sequence 60, Application US/10498788  
; Publication No. US20050118594A1  
; GENERAL INFORMATION:  
; APPLICANT: INCYTE GENOMICS, INC.  
; APPLICANT: Chawla, Narinder K.  
; APPLICANT: Lee, Soo Yeun  
; APPLICANT: Ring, Huijun Z.  
; APPLICANT: Lee, Ernestine A.  
; APPLICANT: Forsythe, Ian J.  
; APPLICANT: Khare, Reena  
; APPLICANT: Tran, Uyen K.  
; APPLICANT: Kabie, Amy E.  
; APPLICANT: Richardson, Thomas W.  
; APPLICANT: Emerling, Brooke M.  
; APPLICANT: Lindquist, Erika A.  
; APPLICANT: Baughn, Mariah R.  
; APPLICANT: Hafalia, April J. A.  
; APPLICANT: Jin, Pei  
; APPLICANT: Swarnakar, Anita  
; APPLICANT: Li, Joana X.  
; APPLICANT: Marquis, Joseph P.  
; APPLICANT: Lee, Sally  
; APPLICANT: Gorvad, Ann E.  
; APPLICANT: Sprague, William W.  
; APPLICANT: Becha, Shanya D.  
; APPLICANT: Elliott, Vicki S.  
; TITLE OF INVENTION: ENZYMS  
; FILE REFERENCE: PF-1312 PCT  
; CURRENT APPLICATION NUMBER: US/10/498,788  
; CURRENT FILING DATE: 2004-06-14  
; PRIOR APPLICATION NUMBER: US 60/340,357  
; PRIOR FILING DATE: 2001-12-14  
; PRIOR APPLICATION NUMBER: US 60/342,962  
; PRIOR FILING DATE: 2001-12-20  
; PRIOR APPLICATION NUMBER: US 60/343,558  
; PRIOR FILING DATE: 2001-12-21  
; PRIOR APPLICATION NUMBER: US 60/351,107  
; PRIOR FILING DATE: 2002-01-22  
; NUMBER OF SEQ ID NOS: 84

SOFTWARE: PERL Program  
; SEQ ID NO 60  
; LENGTH: 2352  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc\_feature  
; OTHER INFORMATION: Incyte ID No: 7506139CB1  
US-10-498-788-60

Query Match 78.08; Score 1318.4; DB 21; Length 2352;

Best Local Similarity 89.38; Pred. No. 0;  
Matches 1509; Conservative 0; Mismatches 1; Indels 180; Gaps 1;

QY	1	GC	CGCCCTTACAGAGACCCCGAGGACGAGTGTCTTCCAGGGCAAGAAAGCATCCC	60
DB	63	GC	CGCCCTTACAGAGACCCCGAGGACGAGTGTCTTCCAGGGCAAGAAAGCATCCC	122
QY	61	CC	GATCACAGTGACCGCTTCTGATCAGAGTGGAGGATCGTGAATGACGACGATC	120
DB	123	CC	GATCACAGTGACCGCTTCTGATCAGAGTGGAGGATCGTGAATGACGACGATC	182
QY	121	CT	TTTAGCGTGATGTCAGCTGAGATGCTTGTATTAACCAATCGGAGAAACCTCAT	180
DB	183	CT	TTTAGCGTGATGTCAGCTGAGATGCTTGTATTAACCAATCGGAGAAACCTCAT	242
QY	181	CG	TCCCTGGGGGATCAAGACCATGACGCGCTGATGGTCTTCTGTTGGCGT	240
DB	243	CG	TCCCTGGGGGATCAAGACCATGACGCGCTGATGGTCTTCTGTTGGCGT	302
QY	241	TG	AGTCCACAACAAGCTGCAGATGCTCTGGGCAATGACACCGCTGACGACTTCTG	300
DB	303	TG	AGTCCACAACAAGCTGCAGATGCTCTGGGCAATGACACCGCTGACGACTTCTG	362
QY	301	TC	AGGACCAAGGACGCTAGCAGGAGGAGCAACCATCATCTTGGACCACTTCTCC	360
DB	363	TC	AGGACCAAGGACGCTAGCAGGAGGAGCAACCATCATCTTGGACCACTTCTCC	422
QY	361	CG	ACGGGTGTGAGCTGTGGCGCTACGAGCAGTGGCGAGGGCGGACGACGCG	420
DB	423	CG	ACGGGTGTGAGCTGTGGCGCTACGAGCAGTGGCGAGGGCGGACGACGCG	482
QY	421	GG	CTGTGCGACTACTCTCTGCACTGAGCATCACCCGATGGCATGAGAGCATCAAGGA	480
DB	483	GG	CTGTGCGACTACTCTCTGCACTGAGCATCACCCGATGGCATGAGAGCATCAAGGA	542
QY	481	GG	AGTGGAGGCTTGTCTAAGGAGAGGGTGTGAATCTTCTGTTCTTCTATGGCATTA	540
DB	543	GG	AGTGGAGGCTTGTCTAAGGAGAGGGTGTGAATCTTCTGTTCTTCTATGGCATTA	602
QY	541	CA	AGGACCGGTGCGAGTGCAGCAGCAGATGTACGAGATCTTACGATCATCCGGGA	600
DB	603	CA	AGGACCGGTGCGAGTGCAGCAGCAGATGTACGAGATCTTACGATCATCCGGGA	662
QY	601	CT	TGGGGGCTTGGCCCGAGGTGCACTGAGAACGGGACATCGTGGAGGAGGACAGAA	660
DB	663	CT	TGGGGGCTTGGCCCGAGGTGCACTGAGAACGGGACATCGTGGAGGAGGACAGAA	722
QY	661	GC	GTTGCTGGAGTCTGGGATCACTGCGCCCGAGGGCCACGTGCTCAGCCCGAGGA	720
DB	723	GC	GTTGCTGGAGTCTGGGATCACTGCGCCCGAGGGCCACGTGCTCAGCCCGAGGA	782
QY	721	GG	TGGAGGCTGAGCGGTGTACCGAGCTGTACCATCGCCCAAGCAGCAACTGCCGCT	780
DB	783	GG	TGGAGGCTGAGCGGTGTACCGAGCTGTACCATCGCCCAAGCAGCAACTGCCGCT	842
QY	781	GT	AGCTCACCAAGGTGATGAGCAAGGGGGCGGCGCCATCGCTCAGGCCAAGCGCAG	840
DB	843	GT	AGCTCACCAAGGTGATGAGCAAGGGGGCGGCGCCATCGCTCAGGCCAAGCGCAG	902
QY	841	AG	GGTGGTGTCTGTTGGGAGGCCATCACCGCAGCTGGGCAACCGGTTCACTTA	900
DB	903	AG	GGTGGTGTCTGTTGGGAGGCCATCACCGCAGCTGGGCAACCGGTTCACTTA	962

## RESULT 7

US-10-948-947A-48  
; Sequence 48, Application US/10948947A  
; Publication No. US20050130924A1  
; GENERAL INFORMATION:  
; APPLICANT: Monia, Brett P.  
; APPLICANT: Freier, Susan M.  
; APPLICANT: Manoharan, Muthiah  
; APPLICANT: Gaarde, William A.  
; APPLICANT: Griffey, Richard H.  
; APPLICANT: Swayze, Eric E.  
; APPLICANT: Bennett, C. Frank  
; TITLE OF INVENTION: ANTISENSE INHIBITION VIA RNASE H-INDEPENDENT REDUCTION IN mRNA  
; FILE REFERENCE: ISPH-0871  
; CURRENT APPLICATION NUMBER: US/10/948, 947A

/ CURRENT FILING DATE: 2004-09-24  
 / PRIOR APPLICATION NUMBER: 60/392,020  
 / PRIOR FILING DATE: 2002-06-26  
 / PRIOR APPLICATION NUMBER: 10/461,163  
 / PRIOR FILING DATE: 2003-06-13  
 / NUMBER OF SEQ ID NOS: 115  
 / SOFTWARE: FastSeq for Windows Version 4.0  
 / SEQ ID NO: 48  
 / LENGTH: 2947  
 / TYPE: DNA  
 / ORGANISM: H. sapiens  
 / FEATURE:  
 / NAME/KEY: CDS  
 / LOCATION: (178)...(1896)  
 / OTHER INFORMATION: antisense oligonucleotide  
 / US-10-948-947A-48

Query Match 58.0%; Score 980.2; DB 22; Length 2947;  
 Best Local Similarity 74.1%; Pred. No. 2.2e-266;  
 Matches 1240; Conservative 0; Mismatches 433; Indels 0; Gaps 0;

QY	18	CCCCAGGAGCAGGATGTCTTCCAGGCGCAAGAAAGCATCCCCCGGATCACGAGTACC	77
DB	164	CGCCCCCGGAGAGATGCTTATCAGGGGAAGAAATAATTCCACGCATCAGCAGCGATC	223
QY	78	GCCTTCTGATCAGAGTGGGAGGATCGTGAATGACGACCACTCTTTTACGCTGATGTGC	137
DB	224	GTCTTCTGATCAAAAGTGGCAAGATGTGAATGATGACCACTCTTCTATGCGACATAT	283
QY	138	ACGTGGAAGATGCTTGTATAAAACAAATCGGAGAAACCTCATCTGCTCCCTGGGGGCATCA	197
DB	284	ACATGGAGATGGTTGATCAACCAATAGGAGAAACCTGTATGTGCCAGAGGGGTGA	343
QY	198	AGACCATGTAGCGCCACGCGCTGATGGTCTTCTGGTGGCGTGTGACGTCCACAAAGGC	257
DB	344	AGACCATCGAAGCCCACTCCAGAATGGTATCCCTGGAGGAAATTGACGTGCACACTCGCT	403
QY	258	TGCAGATGCTGTCTGGGATGACACCGGTGACGACTTCTGTCCAGGGGACCAAGGCAG	317
DB	404	TCAGATGTCCAGACGAGGGGAATACATCAGCTGATGACTTCTTCCAGGGGAACCAAGGCAG	463
QY	318	CGCTAGCAGGAGAAACCAACATGATCTTGGACCACTCTTCCCGGACACAGGGTGTGAGCC	377
DB	464	CCCTGGCCGAGGAAACCAACATGATCATCGACCATGTGTCTGTAGCCCGGACAGCC	523
QY	378	TGCTGGCGCTACGACAGTGGCGGAGCGCGGACAGCGCGGCTGTCTGCGACTACT	437
DB	524	TATTGGCAGCCTTTTGATCAGTGGAGGAGTGGCGGACAGCAAGTCTCTGTGTACTATT	583
QY	438	CCCTGCACGTGGACATCACCCGATGGCATGAGAGCATCAAGGAGGAGCTGGAGCCCTGG	497
DB	584	CGCTGCACGTGGACATCACCGAGTGGCAACAGGGCATCCAGGAGGAGATGGAAGCTCTGG	643
QY	498	TCAAGAGAGAGGTGTGAATCTCTCTGTGTCTTCTATGGCATACAAGGACCGGTGCCAGT	557
DB	644	TGAAGAACCCAGGGGTAAACTCTCTCTGTGTACATGGCTTTCAAGATCGGTTCCAGC	703
QY	558	GCAGCGACAGCCAGATGTACGAGATCTTCAAGCATCATCCGGACCTGGGGGCTTGGCCC	617
DB	704	TGACGGATTTCCAGATCTATGAAGTACTGAGCGGTGATCCGGGATATTGGTGCCATAGCTC	763
QY	618	AGGTGACGCTGAGAACGGGGACATCGTGGAGGAGGAGCAGAGCGGTCTCTGAGCTCG	677
DB	764	AGTCCATGACAGAAATGTGTGATCATCATTCGACGAGAAACAGCAGAGGATCCTGGATCTGG	823
QY	678	GCATCACTGCCCCGAGGGCCAGTGTCTCAGCCACCCCGGAGGAGGTGGAGGCTGAGCGCG	737
DB	824	GCATCAGAGCCCCGAGGGGACAGTGTCTGAGCCGGCCAGAGGAGTTCGAGGCTGAAGCTG	883
QY	738	TGTACCGAGTGTCCATCGCCACAGCAGGCAAACTGCCCGCTGTACGTCCACCAAGGTGA	797
DB	884	TGAACCGGTTCATCACCATTTGCCAATCAGACCAACTGCCCGCTGTATGTCCACCAAGGTGA	943

QY	798	TGAGCAAGGGGGCGCGCGCAGCGCATCGCTCAGGCCAAGCGCAGAGGGGTGCTGCTTTG	857
DB	944	TGAGCAAGAGTGTCTGCTGAAGTCAATGCCCGCAGCAGGAAAGGAACTGTGTTGATG	1003
QY	858	GGGAGCCCATCACCGCCAGCCTGGGACCGCAGCGTTTCACTACTGGAGCAAGAACTGGG	917
DB	1004	GTGAGCCCATCACTGCCAGCCTGGGACTGATGGCTCTCATTTATGGCAAGAACTGGG	1063
QY	918	CCAAGGCTGACGCTTCTGTCATCATACCCCTCTGTCAACCCAGACCCCAACCGGAGACC	977
DB	1064	CCAAGGCGCTGCTTTGTGCTCTCCACCTTTCAGCCCTTTCAGCCCGCAGCACTCCAGACT	1123
QY	978	ACCTCACCTGCTTGTCTCCAGCGGGACCTCCAGGTGACGACGCGCCCACTGCACCT	1037
DB	1124	TTCTCAACTCGTGTGCTCTGTGGAGACTCCAGGTCACTGGAGTGGCCACTGTACCT	1183
QY	1038	TCACCACTGCCCAGAGGCTGTGGGCAAGGACAACTTCCCGCTGATCCCGGAGGACCA	1097
DB	1184	TCACCACTGCCCAGAGGCTGTGGGCAAGGATACTTCACTTGTATTCAGAGGGCACCA	1243
QY	1098	AGGCAATTGAGGAGCGCATGTGATGCTGGAGAAATGTGTGGCTCTGGGAAGATGG	1157
DB	1244	ATGGCACTGAGGAGCGGATGTCTCATTTGGGATAAAGCTGTGGTCACTGGGAAGATGG	1303
QY	1158	ACGAGAACTGAGTTGCTGCGGTGACAGTACAAATGCTGCCAAATCTTCAATTTTACC	1217
DB	1304	ACGAGAACAGATTTGTGGCTGTGACTAGCACCAACGACGCCAAAGTCTTCAATCTTACC	1363
QY	1218	CAAGGAAGGGGCGAGTGGCTGTGGGCTGTGACGCTGACCTGGTTCATATGGAACCCCAAGG	1277
DB	1364	CACGGAAGGTCGTATCTCGTGGGATCTGACGACACCTGGTGTCTGGGACCTTGACA	1423
QY	1278	CCACCAAGATCATCTCTGCCAAGACCAACAATCTGAACTGGAGTACAACTTTCGAGG	1337
DB	1424	GTGTGAAGACCATCTCTGCCAAGACGACACAAGTCTTGAAGTACAACTTTTGAAG	1483
QY	1338	GAGTGAAGTGGCGGGAGCGCTGCGGTGTCTAATAGTCAAGGCGGAGTGGCGCTGGAGG	1397
DB	1484	GAATGAGTGTGGGGCTCCCACTGGTGTATCAGCCAGGGCAAGATTGTCTGGAGG	1543
QY	1398	ACGGGAAGATGTTTGTACCCCGGGGGCGGCTCTGCTCTCGTCCCTCGGAAAACTTCCCGG	1457
DB	1544	ACGGCAGCTTGATGTACGGAAGGCTCAGGACGCTTACATTCCTCCGGAAGCCCTTCCCTG	1603
QY	1458	ACTTTGTCTAAGAGGATCAAGTCTGCAACAGGCTGGGGAGATCCAGGTGTGCGCCC	1517
DB	1604	ACTTTGTGTACAAACGATCAAGGCAAGGAGCAGGCTGGCTGAGCTGAGGGGGTCCCTC	1663
QY	1518	GTGGGCTGTACAGCGGCGGCTCCAGAGTGTGTGCTGCTGCAAGCCAGGGAGTGGCG	1577
DB	1664	GTGGCTGTATGATGAGCCCGTATGCGAGTGTCTGTGACGCCCAAGACGTTCACTCCGG	1723
QY	1578	CTCCGCGCGCGCTCTGCTCCAGGCAAGATCTCTCGTGTCTCTGTGCGCAACTTACATC	1637
DB	1724	CCTCATCAGTAAAGACATCCCTTCCCAAGCAGCAGCGCCACCTGTTTGGAACTTCACC	1783
QY	1638	AGTGGGGTTCAGCCTTATCTGGGTCTCAGGTGATGACCACTGCGCCCGGAGCC	1690
DB	1784	AGTCTGTTTTCAGCTTGTCTGTGTGCTCAGATTGACGACAACTTCCCGCGCGC	1836

RESULT 8  
 US-09-986-632-3  
 ; Sequence 3, Application US/09986632  
 ; Patent No. US2002011944A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: AGUERA, Michelle  
 ; TITLE OF INVENTION: Modulation of Ulip/CRMP activity for the prevention or  
 ; treatment of myelin disorders  
 ; FILE OF INVENTION: P06974501/BAS  
 ; CURRENT APPLICATION NUMBER: US/09/986,632  
 ; CURRENT FILING DATE: 2001-11-09  
 ; PRIOR APPLICATION NUMBER: US 60/246,751

		PRIORITY FILING DATE: 2000-11-09		57.1%; Score 965.4; DB 9; Length 1829;			
		NUMBER OF SEQ ID NOS: 30		Pred. No. 3.1e-262;			
		SOFTWARE: PatentIn Ver. 2.1		0; Mismatches 441; Indels 0; Gaps 0;			
		SEQ ID NO 3		Conservative			
		LENGTH: 1829					
		TYPE: DNA					
		ORGANISM: Homo sapiens					
		US-09-986-632-3					
		Query Match					
		Best Local Similarity					
		Matches 1230;					
QY	20	CCGAGGAGGAGTCTTCCAGGGCAGAAAAGCATCCCCGGGATCAGAGTGACCGC	79				
DB	60	CCGAGGAGGAGATGTCCTATCAGGGGAGAAAATATCCACGCTACGAGCGATCGT	119				
QY	80	CTTCTGATCAGAGTGGGAGATCGTAATGACGACGAGTCTTTTACGCTGATGTCAC	139				
DB	120	CTTCTGATCAAGGAGGTAATATGTTAATGATGATGATGATGATGATGATGATGAT	179				
QY	140	GTGGAAGATGGCTTGTATATAAACAATCGAGAAAACCTCATCTGCTGGGGGCATCAAG	199				
DB	180	ATGGAAGATGGTGTATCAAGCAATAGGAGAAAATCTGATTTGCGAGGAGGATGAAG	239				
QY	200	ACCAATTGACGCCACACGCTGATGTCCTTCTGCTGGGCGTTGACGCTGCCACACAGGCTG	259				
DB	240	ACCATCGAGGCCCACTCCCGGATGATGATCCCGGAGGAATTCACGCTCCACACTCGTTTC	299				
QY	260	CAGATGCTGCTCTGGGATGACACACGCTGACGATCTTCTGTCAGGGGACCAAGGACGG	319				
DB	300	CAGATGCTGATCAGGGAATGACGCTGCTGATGATGATTTCTTCAAGGAACCAAGGGCGCC	359				
QY	320	CTAGCAGGAGGAACCACTATGCTTGGACCACTCTTCCCGACACACGCGTGTGACCTG	379				
DB	360	CTGGCTGGGGGAACCACTATGATGATGATGATGATGATGATGATGATGATGATGATG	419				
QY	380	CTGGCGGCTTACGAGCAGTGGCGGAGCGGGCGGACAGCGCGCTCTGTCGACTACTCC	439				
DB	420	CTGCTGCTCTTACGAGTGGGGAATGGCGCGACAGCAGCAAGTCTCTGTGACTACTCT	479				
QY	440	CTGCACTGAGCATACCCGATGGCATGAGAGCATCAAGAGAGCTGAGGGCCCTGGTC	499				
DB	480	CTGCATGTGGACATCAGCGAGTGGCATTAAGGGCATCCAGGAGGATGGAAGCGCTTGTG	539				
QY	500	AAGGAGAGGGTGTGAACTCTTCTGCTCTTCAATGCAATACAGAGACGCTGCAAGTGC	559				
DB	540	AAGGATCAAGGGTAAATCTTCTCTGCTGTAATGGCTTTCAAGATGCTTCCAGCTA	599				
QY	560	AGCGACAGCCAGATGTACGAGATCTTTCAGCATCATCCGGGACCTGGGGGCTTGGGCCAG	619				
DB	600	ACGGATTGCCAGATTTATGAAGTACTGAGTGTGATCGGGATATTGGGCCATAGCCCAA	659				
QY	620	GTGCACTGTAGAACGGGGACATCTGTGGAGGAGGAGAGAGCGGTTGTGGAGCTCGGC	679				
DB	660	GTCCACGCGAANAATGCGACATCATTTGACAGGAGGAGCAGCAGAGGATCTGGATCTGGGC	719				
QY	680	ATCACTGGCCCCGAGGSCACGCTGCTCAGCCACCCCGAGGAGGTGAGGCTGAGGCGGTG	739				
DB	720	ATCACGGGCCCCGAGGAGCATGTGCTGAGCCGACCTGAGGAGGTCCAGGCGCGAAGCCGTG	779				
QY	740	TACCGAGCTGTCAACATCGCAAGCAGGCAAACTGCGCGCTGTCACCTCAAGGTGATG	799				
DB	780	AATCGTGCCATCAACATCGCCACACAGCACTGCGCGCTGTATATCACCAGGTGATG	839				
QY	800	AGCAAGGGGGCGGCCGACGCTATGCTCAGGCCAACGCGAGAGGGGTGCTGCTGTTTGGG	859				
DB	840	AGCAAAAGCTCTGCTGAGGTATCTGCGCCAGGCAAGGAAAGAACTGTGGTGTATGCG	899				
QY	860	GAGCCCATCAACCGCCAGCTGGGACCGAGCGGTTCACACTCTGAGGACGAAGCTGGGCC	919				
DB	900	GAGCCCATCACTGCCAGCTGGGAACGGACCGGCTCCCTCCTTCTGAGGACGAAGCTGGGCC	959				

RESULT 9  
US-10-062-674-1971  
; Sequence 1971, Application US/10062674  
; Publication No. US20040005559A1  
; GENERAL INFORMATION:  
; APPLICANT: Loring, Jeanne F.; Kaber, Matthew R.  
; TITLE OF INVENTION: MARKERS OF NEURONAL DIFFERENTIATION AND MORPHOGENESIS  
; FILE REFERENCE: PA-0026-1 CIP  
; CURRENT APPLICATION NUMBER: US/10/062,674  
; CURRENT FILING DATE: 2002-01-30  
; PRIOR APPLICATION NUMBER: US 09/625,102  
; PRIOR FILING DATE: 2000-07-24  
; NUMBER OF SEQ ID NOS: 2217  
; SOFTWARE: PERL Program  
; SEQ ID NO 1971  
; LENGTH: 4435  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc\_feature

OTHER INFORMATION: Incyte ID No. US20040005559A1 371288.7  
US-10-062-674-1971

Query Match 57.1%; Score 965.4; DB 17; Length 4435;  
Best Local Similarity 73.6%; Pred. No. 3.7e-262;  
Matches 1230; Conservative 0; Mismatches 441; Indels 0; Gaps 0;

```
QY 20 CCAGAGCAGGATGCTTCCAGGCAAGAAAGCATCCCCGGATCCAGAGTGACCG 79
DB 229 CCAGAGAGAGATGCTTATCAGGGGAAGAAAATATTCACGCATCAGAGCGATCGT 288
QY 80 CTTCTGATCAGAGTGCGAGATCGTGAATGACGACAGTCCCTTTTACGCTGATGTCAC 139
DB 289 CTTCTGATCAAGGAGGTAAATGTTAATGATGACCAAGTCGTTCTATGACAGATATAC 348
QY 140 GTGGAAGATGGCTTGATAAACAATTCGGGAAGAACTCATGCTCCCTGGGGGCATCAAG 199
DB 349 ATGGAAGATGGGTTGATCAAGCAAAATAGGAGAAAATCTGATTTGTCCAGGAGGATGAAG 408
QY 200 ACCATTGACGCCACCGGCTGATGTCCTTCTGCTGGGTTGACGTCACACAAGGCTG 259
DB 409 ACCATGAGGCCCACTCCCGATGGTGATCCCCGGAGGAATTGACGTCCACACTCGTTTC 468
QY 260 CAGATGCTGCTCCTGGGCAATGACACCGGCTGACGACTTCTGTCAAGGCAACCAAGGACGG 319
DB 469 CAGATGCTGATCAGGGAATGACGCTCTGCTGATGATTTCTTCAAGGAACCAAGGCGGCC 528
QY 320 CTAGCAGGAGGAACCAACCATGATCTTTGGACCAAGTCTTCCCGACAGCGGTGTGACGCTG 379
DB 529 CTGGCTGGGGGAACCACTATGATCAATTGACCAAGCTTGTCTGAGCTGGGCAAGCCTG 588
QY 380 CTGGCGGCTTACAGAGTGGGGGAGCGGGCGGACAGCGGCTCTGCTCGACTACTCC 439
DB 589 CTGCTGCTCTTACCAAGTGGAGGAATGGGCGGACAGCAAGTCTCTGCTGTGACTACTCT 648
QY 440 CTGCAAGTGGACATCAACCGATGGCATGAGAGCATCAAGGAGAGCTGGAGGCGCTTGCTC 499
DB 649 CTGATGTGGACATCAGCGAGTGGCATAGGGCATCCAGGAGAGATGGAGGCTTTGTG 708
QY 500 AAGGAGAGGTGTGAATCTCTCTCTGCTGCTTCAAGGATCAAGGACCGGTGCCAGTGC 559
DB 709 AAGGATCAAGGGTAAATCTCTCTCTGCTGCTATGCTGCTTCAAGATGCTTCCAGCTA 768
QY 560 AGCAGAGCGGATGTACGAGATCTTCAGCATCATCGGACCTGGGGGCTTTGGCCAG 619
DB 769 ACGATGCCAGATTTATGAAGTACTGAGTGTATCCGGATATTTGGCGCCATAGGCCAA 828
QY 620 GTGCAAGCTGAGAACGGGACATCTGTGGAGGAGCAGAAAGCGTTGTCTGGAGCTCGGC 679
DB 829 GTCCACGCAGAAAATGGCGACATCATTCAGAGAGGACGACAGAGATCCTGGATCTGGGC 888
QY 680 ATCACTGGCCCCGAGGCGCACGTGCTCAGCCACCCGAGAGAGTGAGGCTGAGGCGGTG 739
DB 889 ATCAAGGCCCCGAGGAGACATGTGCTGAGCGGACCTGAGGAGGTGAGGCGGAGCGGTG 948
QY 740 TACCGAGCTGTACCATGCGCAGCAGGCAAACTGCGCGTGTACGTCAACAGGTGATG 799
DB 949 AATGCTGCCATCACTACCATGCGCAACCAAGACCAATGCGCCGCTGTATATCAACAGGTGATG 1008
QY 800 AGCAAGGGGCGGCGCAGCGCATCTGTCCAGCCAAAGCGCAGAGGGGTGGTGTGTTTGGG 859
DB 1009 AGCAAAAGCTCTCTGAGGTGATCGCCCGAGCAGGAAGGGAATGTGTGTATGGC 1068
QY 860 GAGCCCATCAACCGCAGCTGGGACCGACGCTTCACTACTGTGAGCAAGAACTGGGCC 919
DB 1069 GAGCCCATCACTGCCAGCTTGGGAAAGGACGCTCCCATTTACTTGGAGCAAGAACTGGGCC 1128
QY 920 AAGGCTGACGCTTCTGATCATCACTCCCTGTCAACCCAGACCCACCGAGGAGCCAC 979
DB 1129 AAGGCTGCTGCTTGTGCTCTCCCACTCCCACTTGGAGCCCTGATCCAACTCCAGACTTT 1188
QY 980 CTCACTGCTGTGCTCAGCGGGGACCTCCAGGTGACAGGACGCGCCACTGCACTTC 1039
```

```
DB 1189 CTCAACTCTTGCTGTCTGTGGAGACCTCCAGGTCAAGGAGTCCCACTTTCACGCTTT 1248
QY 1040 ACCACTGCCAGAAAGGCTGTGGGCAAGGACAACTTGGGCTGATCCCCGAGGGCAAC 1099
DB 1249 AACACTGCCAGAAAGGCTGTAGGAAAGGACAACTTCACTCTGATTCGAGGGGACCAAT 1308
QY 1100 GGCATTGAAGAGCGGATGTGATGTCTGGGAGAAATGTGTGGCTCTCTGGGAAGATGGAC 1159
DB 1309 GGCATGTAGAGAGCGGATGTCCGTCTCTGGGACAGGCTGTGTCTCTGGGAAGATGGAT 1368
QY 1160 GGAATGAGTTCGTCCGGGTGACCAAGTACAAATGTGCAAAATCTTCAATTTTACCCA 1219
DB 1369 GGAACCAAGTTCGTGGCTGTGACAGCAACCAATGAGCCAAAGTCTTCAACTTTACCCG 1428
QY 1220 AGGAAGGGGAGTGGCTGTGGGCTGTGAGCTGACCTGACCTGATATGGAACCCCAAGGCC 1279
DB 1429 CGGAAGGGCGCATGCTGTGGGATCCGATCCGACCTGCTCATCTGGGACCCCGACAGC 1488
QY 1280 ACCAAGATCATCTGCGCAAGACCAACATCTGAACGTGGAGTACAACTCTTCGAGGGA 1339
DB 1489 GTTAAACCATCTTGGCAAGACACAAACAGCTCTCTCGAGTACAACTCTTGAAGGC 1548
QY 1340 GTGGAGTCCGGGAGCGCTCCGCTGTGTCATAAGTCAAGGCGGAGTGGCTGGAGGAC 1399
DB 1549 ATGGAGTCCGGGCTCCCACTGGTGTGTCATCAGCCAGGGGAAGATTGTCTCTGGAGGAC 1608
QY 1400 GGAAGATGTTGTCAACCCGGGGGGGCGGCTGCTGCTCGGAAACATTCCTCCGAC 1459
DB 1609 GGCACCTGCTATGTCACCGAAGGCTCTGGACGCTACATTCCTCCGGAAGCCCTTCCCTGAT 1668
QY 1460 TTTGTCTCAAGAGGATCAAGCTCGCAACAGCTGGCGGAGATCCACGCTGTGCCCCCGT 1519
DB 1669 TTTGTTTACAGGTATCAAGGCAAGGACAGGCTGGCTGAGCTGAGAGGGGTTCCTCGT 1728
QY 1520 GGGCTGTATGACGGGCGCTCCACAGGTGATGTGTGCTGCCAAGCAGGAGTGGCGCT 1579
DB 1729 GGCCTGTATGACGAGCCGCTGTGTAAGTGTCTGTGACGCCCAAGACAGTCACTCCAGCC 1788
QY 1580 CCGGCCGCGGCTCTGCGCCAGGCAAGATCTCGGTGCTCTGTGGGCAACCTTACATCAG 1639
DB 1789 TCTCGCCCAAGACGCTCTCTGCAAGCAGGCGGCCACCTGTCTCGGAACCTTGCACAG 1848
QY 1640 TCGGGGTTCAGGCTATCTGGGCTCTCAGGCTGATGACCATGCGCCCGACGCGC 1690
DB 1849 TCTGGATTGATTTGTCTGGTCTCAGATTGATGACAACTTCCCCGCGCGC 1899
```

## RESULT 10

```
US-10-175-523-140
; Sequence 140, Application US/10175523
; Publication No. US20030096264A1
; GENERAL INFORMATION:
; APPLICANT: Brockman, Jeffrey
; APPLICANT: Evans, David
; APPLICANT: Hook, Derek
; APPLICANT: Klimczak, Leszek
; APPLICANT: Laeng, Pascal
; APPLICANT: Palfreyman, Michael
; APPLICANT: Rajan, Prithi
; TITLE OF INVENTION: MULTI-PARAMETER HIGH THROUGHPUT SCREENING ASSAYS (MPHTS)
; FILE REFERENCE: 3235/10795-US3
; CURRENT APPLICATION NUMBER: US/10/175,523
; CURRENT FILING DATE: 2002-06-18
; PRIOR APPLICATION NUMBER: US 60/299,151
; PRIOR FILING DATE: 2001-06-18
; PRIOR APPLICATION NUMBER: US 60/317,828
; PRIOR FILING DATE: 2001-09-07
; PRIOR APPLICATION NUMBER: US 60/325,150
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: US 60/333,047
; PRIOR FILING DATE: 2001-11-14
; PRIOR APPLICATION NUMBER: US 60/349,936
; PRIOR FILING DATE: 2002-01-18
```

;; PRIOR APPLICATION NUMBER: US 60/361,834  
;; PRIOR FILING DATE: 2002-03-04  
;; NUMBER OF SEQ ID NOS: 197  
;; SOFTWARE: PatentIn version 3.1  
;; SEQ ID NO 140  
;; LENGTH: 4459  
;; TYPE: DNA  
;; ORGANISM: Homo sapiens  
US-10-175-523-140

Query Match 57.0%; Score 963.8; DB 14; Length 4459;  
Best Local Similarity 73.5%; Pred. No. 1.1e-261;  
Matches 1229; Conservative 0; Mismatches 442; Indels 0; Gaps 0;

Qy	20	CCAGGAGCAGGATGTCCTTCCAGGGCAAGAAAGCATCCCGGATCACGAGTGACCGC	79
Db	263	CCAGGAGAGAGATGTCCTATCAGGGGAGAGAAATATCCACGATCACGAGCGATCGT	322
Qy	80	CTTCTGATCAGAGTGGGAGGATCGTGAATGACGACGAGTCCCTTTAGCGTGTATGTCAC	139
Db	323	CTTCTGATCAAGGAGGTAAATTTGTTAATGATGACCAAGTCGTTCTATGCGACATATAC	382
Qy	140	GTGAGATGCTGTGATTAACAAATCGGAGAAACCTCATCGTCCCTGGGGGCATCAAG	199
Db	383	ATGGAAGATGGTTGATCAAGCAATAGGAGAAATCTGATTTGCCAGGAGGTGAAG	442
Qy	200	ACCATGACGCCACGCGCTGATGCTCTTCTGCTGGCGTTGACGTCCACACAAAGGCTG	259
Db	443	ACCATCGAGGCCACTCCCGGATGTTGATCCCCGAGGAATTAAGTTCACACTCGTTTC	502
Qy	260	CAGATGCTGTCTGGGATGACACCGGCTGACGATCTTCTGTCAGGGCAACAAAGGACGG	319
Db	503	CAGATGCTGTATCAGGGAATGACGCTGCTGATGATTTCTTCAAGGAACCAAGGCGCC	562
Qy	320	CTAGCAGAGGAAACCATGATCTTGGAACAGTCTTCCCGACAGCGGTGTGAGCGTG	379
Db	563	CTGGCTGGGGGAACCATATGATCATTTGACACAGTCTTCTGAGCCTGGGACAGCGTG	622
Qy	380	CTGGCGCCTACGAGCAGTGGCGGAGCGGCGGACAGCGCGCTCTCGACTACTCC	439
Db	623	CTCGCTGCTTTCAGCTAGTGGAGGATGGCGGACAGCAAGTCTCTGTGTACTCT	682
Qy	440	CTGCACTGTGACATCACCCGATGGCATGAGAGCATCAAGAGGAGTGGAGGCGCTGGTC	499
Db	683	CTGCATGTGACATCAGCGAGTGGCAATAAGGCAATCAGGAGGAGATGGAAGCGCTGTG	742
Qy	500	AAGGAGAGGTGTGAACTCTTCTCTGCTTTCATGTCATACAGGACCGGTGCCAGTGC	559
Db	743	AAGGATCACGGGGTAAATCTTCTCTGTGTACATGCTTTCAAGAGATCGCTTCCAGCTA	802
Qy	560	AGCGACAGCCAGATGTACGAGATCTTCAGCATCATCCGGACCTGGGGGCTTGGCCGAG	619
Db	803	ACGGATTGCCAGATTTATGAGTACTGATGTGATCCGGGATATTGGCGCCATAGCCCAA	862
Qy	620	GTGCACGCTGAGAACGGGGAATCGTGGAGGAGGAGCAGAACGGTTGCTGGAGCTCGGC	679
Db	863	GTCCACGACAGAAATGGCGACATCATTTGACAGAGGAGCAGCAGGATCCTGGATCTGGGC	922
Qy	680	ATCACTGGCCCCGGGCGCACGTGCTCAGCCACCCCGAGGAGTGGAGGCTGAGGGGTG	739
Db	923	ATCACGGGCCCGAGGGACATGTGCTGAGCCGACCTGAGGAGGTGCGAGGCGGAAGCGTG	982
Qy	740	TACCGAGCTGTCAACATCGCCAGCAGGCAAACTGCGCGTGTACGTACCAAGGTGATG	799
Db	983	AATGTGCCATCACCATTCGCCAACGACCACTGCCCCGTGTATATCACCAAGGTGATG	1042
Qy	800	AGCAAGGGGGCGGCGCACGCGCATCGCTCAGGCCAAGCGCAGAGGGGTGGTCTGTTTGGG	859
Db	1043	AGCAAAAGCTCTGCTGAGTCTATCGCCAGGACGGAAGGAAGTGTGTGTATGGC	1102
Qy	860	GAGCCCATACCGCGACGCTGGGACCGGAGGTTCACATCTGAGGACAAAGACTGGGCC	919
Db	1103	GAGCCCATCATCTGCAGCTTGGGAACGGAGCGGTCTCCATTTACTGGAGCAAGAACTGGGCC	1162

Qy	920	AAGGCTGCAGCCTTGTTCATCATCACCCCTGTCAACCCAGACCCACACGCGAGACCAC	979
Db	1163	AAGGCTGTGCTTTGTACCTCCACCCCTTTCAGCCCTGATCCAAACCACTCCAGACTTT	1222
Qy	980	CTCACCCTGCTTGTCTCAGCGGGGAGCTCCAGGTGACAGGCGGCCACTGACACTTC	1039
Db	1223	CTCAACTCCTTGTCTGTGTGAGACCTTCAGGTCAOGGGAGTGGCCATTGACACTTT	1282
Qy	1040	ACCACCTGCCAGAAAGCTGTGGGCAAGGACAACTTCGCGCTGATCCCGAGGGGACCAAC	1099
Db	1283	AACACTGCCAGAAAGCTGTAGGAAGGACAACTTCACCTGATTCGGGAGGACCAAT	1342
Qy	1100	GGCAATTGAGAGGCGCATGTGATGTGTGGAGAAATGTGTGGCTCTGGGAAGATGGAC	1159
Db	1343	GGCACTGAGAGGCGCATTTGCTGTGGGATCCGATCCGACCTGTGTCTCACTGGGAAGATGGAT	1402
Qy	1160	GAGAAATGTTTCTGCGGTGACCATGACAAATGCTGCCAAATCTCAATTTTACCCA	1219
Db	1403	GAGAACAGTTTGTGGCTGTGACCAACCAATGCGAGCCAAAGTCTTCAACCTTTACCCC	1462
Qy	1220	AGGAAGGGGAGTGTGTGGCTCTGACGCTGACCTGTGTCATATGGAACCCCAAGGCC	1279
Db	1463	CGGAAGGGCGCATTTGCTGTGGGATCCGATCCGACCTGTGTCTATCTGGGACCCCGACGC	1522
Qy	1280	ACCAAGATCATCTCTGCCAAGACCCCAATCTGAACGTGGAGTACAACTCTTTCGAGGGA	1339
Db	1523	GTTTAAACCATCTCTGCCAAGACACACACAGCTCTCTCGAGTACAACTCTTTGAAGGC	1582
Qy	1340	GTGGAGTCCCGGGAGCGCTCGCTGTGTGTCATTAAGTCAGGGCGGAGTGGCGCTGGAGGAC	1399
Db	1583	ATGGAGTCCCGCGCTCTCCCACTGTGTGTGTCATCAGCCAGGGGAGGATTTCTCTGGAGGAC	1642
Qy	1400	GGGAAGATGTTTCTCAACCCCGGGGCGGCGCTCTGCTCTCGGAAACATTCGCCGAC	1459
Db	1643	GGCACCTGTATCTACCGAGGCTCTGGACGCTTACATTTCCCGGAGGCCCTTCCCTGAT	1702
Qy	1460	TTTGTCTTACAAGAGGATCAAGCTCGCAACAGGCTGGCGGAGATCCACGCTGTGCCCGCT	1519
Db	1703	TTTGTTTACAAGGTATCAAGGCAAGGACAGGCTGGCTGAGCTGAGAGGGGTTCTCTCGT	1762
Qy	1520	GGGCTGTATGACGGGCGGCTTCCAGAGGTGATGGTGGCTGCCAAGCCAGGAGTGGCGCT	1579
Db	1763	GGGCTGTATGACGAGACTGTGTGTGAAGTGTCTGTGACGCCCAAGACAGTCACTCCAGCC	1822
Qy	1580	CGGGCCGCGGCTCTGCCAGGCAAGATCTCCGTGCTCTCTGCGCAACCTTACATCAG	1639
Db	1823	TCCTCGGCCAAGACGCTCTCTGCCAAGCAGGAGGCCCCACCTGTCCGGAACCTGCACAG	1882
Qy	1640	TCGGGGTTTCAAGCTATCTGGGTCTCAGGCTGATGACCACTCGCCCGGACGC	1690
Db	1883	TCTGGATTCAGTTTGTCTGTGTCTCAGATTGATGACAACTTCCCGCGCGC	1933

RESULT 11

US-10-006-911-11  
; Sequence 11, Application US/10006911  
; Publication No. US20030125274A1  
; GENERAL INFORMATION:  
; APPLICANT: William Gaarde  
; APPLICANT: Andrew T. Watt  
; TITLE OF INVENTION: ANTISENSE MODULATION OF COLLAPLIN RESPONSE MEDIATOR PROTEIN 2 EXP  
; FILE REFERENCE: RTS-0200  
; CURRENT APPLICATION NUMBER: US/10/006,911  
; CURRENT FILING DATE: 2001-11-08  
; NUMBER OF SEQ ID NOS: 89  
; SEQ ID NO 11  
; LENGTH: 4459  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-006-911-11

Query Match

57.0%; Score 963.8; DB 15; Length 4459;



Best Local Similarity 73.5%; Pred. No. 1.1e-261;  
Matches 1229; Conservative 0; Mismatches 442; Indels 0; Gaps 0;

QY 20 CCAGAGCAGGATGTCCTTCAGGCGCAAGAAAAGCATCCCCGGATCAGAGTGACCGC 79  
Db 263 CCAGAGAGAGATGTCCTTATCAGGGGGAAGAAAAATATCCACGCATCAGAGCGATCGT 322

QY 80 CTTCTGATCAGAGGTGGAGATCGTGAATGACGACCCAGTCCCTTTTACGCTGATGTGCAC 139  
Db 323 CTTCTGATCAAGAGGATGAATGTTGATGATGACCAAGTCGTTCTATGACACATATAC 382

QY 140 GTGGAAGATGGTTGATAAAACAAATCGGAGAAAACTCATCTCCTCGGGGCATCAAG 199  
Db 383 ATGGAAGATGGTTGATCAAGCAAAATAGGAGAAAACTGATTTGTGCCAGGAGGTGAAG 442

QY 200 ACCATTGACGCCCAAGCGCTGATGTCCTTCCTGCTGGTGGCGTTGACGTCCACACAAGGCTG 259  
Db 443 ACCATGAGGCCCACTCCCGGATGGTGATCCCGGAGGAATGACGTCCACACTCGCTTTC 502

QY 260 CAGATGCTGCTCGGCGATGACACCGGCTGACGACTTCTGTGAGGSCACCAAGGACGCG 319  
Db 503 CAGATGCTGATCAGGGAATGAGCTGTGCTGATGATTTCTTCAGGAACCAAGGCGGCC 562

QY 320 CTAGCAGGAGGAACCACTATGATCTTTGGACCACTCTTCCCGACACGGGTGTGACCTG 379  
Db 563 CTGCTGGGGGAACCACTATGATCATTTGACCACTGTTGTTCTGAGCTGGGACAAGCCTG 622

QY 380 CTGGCGGCTCAGACGATGGCGGAGCGGGCGGACAGCGCGGCTGCTCGCATCTACTCC 439  
Db 623 CTGCTGCTTTGACCACTGGAGGAATGGGGCGGACAGCAAGTCTCTGCTGACTACTCT 682

QY 440 CTCACGTGACATCACCGGATGCGATGAGAGCATCAAGAGAGAGCTGGAGGCCCTGGTC 499  
Db 683 CTGATGTGACATCAGCGAGTGCGCATAGGGCATCCAGGAGAGATGGAAGCGCTGTG 742

QY 500 AAGGAGAAGGTTGAATCTCTTCTGCTTCTTATGCGCATACAAGGACCGGTGCCAGTGC 559  
Db 743 AAGGATCAGGGGTAAATCTCTTCTCGTGATCATGGCTTTCAGAGATCGTTCAGCTA 802

QY 560 AGGCAGACGATGTACGAGATCTTCAGATCATCTGGGACCTCGGGGCTTGGGCGCAG 619  
Db 803 ACGGATTCGCGAGATTTATGAAGTACTGAGTGTGATCCGGGATATTGGCGCCATAGCCAA 862

QY 620 GTCACGCTGAGAACGGGACATCTGTGAGAGGAGCAGAGCGGTTGCTGGAGCTCGC 679  
Db 863 GTCACGAGAAAATGGCGACATCTTGCAGAGAGCAGCAGAGATCTTGGATCTGGGC 922

QY 680 ATCACTGGCCCCGAGGCCACGTGCTCAGCCACCCGAGAGGTGGAGGCTGAGGGCGTG 739  
Db 923 ATCAGGGCCCCGAGGGACATGTGCTGAGCCGACCTGAGGAGTTCGAGGCCGAAGCGTG 982

QY 740 TACCGAGCTGTACCATCGGCAAGCAGGCAAACTGCCCGCTGTACGTACCAAGGTTGATG 799  
Db 983 AATCGTGCCATCACTACGCAACCAAGACCACTGCGCGCTGTATATCACAAGGTGATG 1042

QY 800 AGCAAGGGGCGCGCAGCCATCGCTCAGGCCCAAGCGAGAGGGTGTGCTGTTTGG 859  
Db 1043 AGCAAGGCTCTGCTGAGGTGATCGGCCAGCGACGGAAGGGAATGTGGTGTATGGC 1102

QY 860 GAGCCCATCAACCGCAGCTGGGCAACGAGCGTTTCACACTACTGAGAGCAAGAACTGGGCC 919  
Db 1103 GAGCCCATCACTGCCAGTTGGGAACGAGCGGCTCCCATCTACTGGAGCAAGAACTGGGCC 1162

QY 920 AAGGCTGACGCTTGTGATCATCACTCCCTGTGCAACCGAGACCCACAGCGGACCAAC 979  
Db 1163 AAGGCTGCTGCTTGTGATCTCCCACTCCCACTGAGCCCTGATCCAACTCCAGCACTTT 1222

QY 980 CTCACCTGCTGTGTCAGCGGGGACCTCAGGTGACAGCGGACGCCACTGCACTTC 1039  
Db 1223 CTCACCTGCTGTGTCAGCGGACCTCAGGTGACAGCGGAGTGGCCCACTTGCAGCTTT 1282

QY 1040 ACCACTGCCAGAGGCTGTGGGCAAGGACAACTTCGCGCTGATCCCCCGAGGCGACCAAC 1099

Db 1283 AACACTGCCAGAGGCTGTAGGAAAAGCAAACTTTCACCCCTGATTCGGAGGGCACCAAT 1342  
QY 1100 GCATTGAGAGGCGCATGTGATGCTCTGGGAGAAATGTGTGCCCTCTCGGAGATGGAC 1159  
Db 1343 GGACTGAGAGCGGATGTCCTGTCATCTGGGACAAAGGCTGTGGTCACTGGGAAGATGGAT 1402

QY 1160 GAGAATGATGCTGCTGCGGTGACCAAGTACAAATGCTGCCAAAATCTTCAATTTTACCCA 1219  
Db 1403 GAGAACCACTGTTGCTGTGACCAAGCACTGAGCCCAAGTCTTCAACTTTTACCCC 1462

QY 1220 AGAAGGGCGAGTGGCTGTGGCTCTGACGCTGACCTGGTCATATGGAACCCCAAGGCC 1279  
Db 1463 CGAAAAGGCGCATGCTGTGGGATCCGATCCGACCTGCTGATCTCTGGACCCCGCAGC 1522

QY 1280 ACCAAGATCATCTGCGCAAGACCCACAACTCTGAAGCTGAGGTACAACTCTTCGAGGGA 1339  
Db 1523 GTTAAAAACCATCTCTGCCAAGACACAAACAGCTCTCTCGGTACAACTCTTTGAAGGC 1582

QY 1340 GTGGAGTCCGGGAGCGCTGCTCGGTGCTATAAGTCAAGGCGCGAGTGGCGCTGGAGGAC 1399  
Db 1583 ATGGAGTCCGGGCTCCCACTGGTGGTCTATCAGCCAGGGGAGATTTGCTGTGAGGAC 1642

QY 1400 GGAAGATGTTGTTCACCCCGGGGGCGGCGCTTGTGCTCCCTCGGAAAAATTCCTCGGAC 1459  
Db 1643 GGCACCTGCTGTCACCGAAGGCTCTGGACGCTACATTCCTCGGAAAGCCCTTCCCTGAT 1702

QY 1460 TTTGTCTACAAGAGGATCAAGCTCGCAACAGGCTGGCGGAGATCCACGGTGTGCCCT 1519  
Db 1703 TTTGTTTACAAAGCTATCAAGGCAAGGACAGGCTGGCTGAGCTGAGAGGGGTTCTCGT 1762

QY 1520 GGGCTGTATGACGGGCGGCTCCACAGGTGATGTTGCTGCCCAAGCCAGGAGTGGCGCT 1579  
Db 1763 GGCCTGTATGACGACCTGTGTGAGTGTCTGTGACGCCCAAGACATCTACTCCAGCC 1822

QY 1580 CCGGCGCGCGCTCTGCCCAAGGCAAGATCTCCGCTCTCTGTGCGAACCTTACTACAG 1639  
Db 1823 TCTCGGCGCAAGACGCTCTCTGCCAAGCAGCAGGCGCCACCTGTCCGAAACCTGCACAG 1882

QY 1640 TCGGGGTTTCAGCTATCTGGGTCTCAGGCTGATGACCACTCGCCGACGC 1690  
Db 1883 TCTGGATTCACTTTGTCTGGTCTCAGATTGATGACAACTTCCCGCGCGC 1933

RESULT 12  
US-10-133-937-8  
; Sequence 8, Application US/10133937  
; Publication No. US20030207278A1  
; GENERAL INFORMATION:  
; APPLICANT: Khan, Javed  
; APPLICANT: Ringner, Markus  
; APPLICANT: Peterson, Carsten  
; APPLICANT: Meltzer, Paul  
; TITLE OF INVENTION: METHODS FOR ANALYZING HIGH DIMENSIONAL DATA FOR CLASSIFYING,  
; TITLE OF INVENTION: DIAGNOSING, PROGNOSTICATING, AND/OR PREDICTING DISEASES AND  
; TITLE OF INVENTION: OTHER BIOLOGICAL STATES  
; FILE REFERENCE: 11613 56US01  
; CURRENT APPLICATION NUMBER: US/10/133,937  
; NUMBER OF SEQ ID NOS: 99  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 8  
; LENGTH: 4459  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-133-937-8

Query Match 57.0%; Score 963.8; DB 17; Length 4459;  
Best Local Similarity 73.5%; Pred. No. 1.1e-261;  
Matches 1229; Conservative 0; Mismatches 442; Indels 0; Gaps 0;

QY 20 CCAGAGCAGGATGTCCTTCCAGGCGCAAGAAAAGCATCCCCGGATCAGAGTGACCGC 79  
Db 263 CCAGAGAGAGATGTCCTTATCAGGGGGAAGAAAAATATTCACGCAATCAGAGCGATCGT 322

QY 80 CTTCTGATCAGAGTGGGAGGATCGTGAATGACGACAGTCTTTTACGCTGTATGTGCAC 139  
Db 323 CTTCTGATCAGAGGAGTGAATTTGTTAATGATGACAGTCTTTTATGACGATATAC 382  
QY 140 GTGGAAGATGGCTGTGATTAACAAATCGGAGAAACCTCATCTGTCCTTGGGGGATCAAG 199  
Db 383 ATGGAAGATGGTGTGATCAAGCAATAGGAGAAATCTGATTTGGCCAGGAGGATGAAG 442  
QY 200 ACCATTGACGCCACGCGCTGATGCTCTTCTGCTGGCTGTGACGTGCCACACAAGGCTG 259  
Db 443 ACCATCGAGGCCCACTCCCGATGGTGTATCCCGAGGAATGACGTCCAACATCTGCTTC 502  
QY 260 CAGATGCTGTCTGCGCATGACACCGGCTGACGACTTCTGTCAAGGACACCAAGGACAGCG 319  
Db 503 CAGATGCTGTATCAGGGAATGACGCTCTGCTGATGATTTCTTCCAAGGAACCAAGGGGCC 562  
QY 320 CTAGCAGAGGAACCAACATGATCTTGGACACAGTCTTCCCGACACAGGGTGTAGGCTG 379  
Db 563 CTGCTGGGGGAACCACTATGATCAITGACACAGTCTTCTGAGCCTGGGACAAAGCTG 622  
QY 380 CTGCGGCTTACGAGCAGTGGCGAGCGGCGGACAGCGGCTGCTGCGACTACTCC 439  
Db 623 CTGCTGCTTTGACCATGAGGGAATGGGCCGACAGCAAGTCTCTCTGTGACTCTCT 682  
QY 440 CTGCACTGGACATCACCCGATGGCATGAGAGCATCAAGGAGGAGCTGGAGGCCCTGGTC 499  
Db 683 CTGATGTGACATCAGCGAGTGGCATAAAGGCAATCAGGAGGAGATGGAAGCGCTTGTG 742  
QY 500 AAGGAGAGGGTGTGAATCTTCTGCTGTCTTATGTCATCAAGGACCGGTGCGAGTGC 559  
Db 743 AAGGATCAGCGGGTAAATCTTCTCTGCTGATCAATGGCTTTCAAGATCGCTTCCAGCTA 802  
QY 560 AGCGACAGCAGATGTACGAGATCTTCAGCATCATCCGAGCATCTGGGGCTTGGCCAG 619  
Db 803 ACGATTCGCGAGTTTATGAGTACTGAGTGTGATCCGGATATTTGGCGCCATAGCCAA 862  
QY 620 GTGCACTGTGAGAACGGGGACATCGTGGAGGAGGAGCAGAACGGTGTGTGAGCTCGGC 679  
Db 863 GTCCACGACAGAAATGGGACATCATTTGACAGGAGGAGCAGAGGATCTTGGATCTGGC 922  
QY 680 ATCACTGGCCGAGGGCCACGTGCTCAGCCACCCCGAGGAGGTGAGGCTGAGGGGTG 739  
Db 923 ATCAGCGGCCCGAGGGACATGTCTGAGCGCACTGAGGAGGTGCGAGGCGCAAGCGGTG 982  
QY 740 TACGAGCTGTACCATCGCCACGACGAGCAAACTGCGCGTGTACGTACCAAGGTGATG 799  
Db 983 AATCGTCCCATCACATCGCCAAACAGACCAACTGCGCGTGTATATCACCAAGGTGATG 1042  
QY 800 AGCAAGGGGGCGGCGGACGATCGCTCAGGCCAAGCGCGCAGAGGGGTGGTCTGTTTGGG 859  
Db 1043 AGCAAAAGCTCTGCTGAGGTCTATCGCCGAGGACGGAAGGAACTGTGTGTATGGC 1102  
QY 860 GAGCCCATCAACGCGACGCTGGGACCGACGAGTGTCACTACTGAGCAAGAACTGGGCC 919  
Db 1103 GAGCCCATCACTGCGAGCTGGGAAACGAGCGGCTCCCACTTACTGAGCAAGAACTGGGCC 1162  
QY 920 AAGGCTCAGCCTTGTGTCATCATCAACCCCTGTCAACCCAGCCCAACGCGGACCCAC 979  
Db 1163 AAGGCTCTGCTTGTGTCATCTCCCACTTGGCCCTTGAAGCCCTGATCCAACCACTCCAGACTT 1222  
QY 980 CTCACTGCTGTCTGCTCAGCGGGGACCTCCAGGTGACAGCAGCGGCCCACTGACCTTC 1039  
Db 1223 CTCACTCTTGTCTGCTGAGGACCTCCAGGTGACGGGCACTGCGCACTTGCACGTTT 1282  
QY 1040 ACCATGCCCAGAGGCTGTGGGCAAGGAACAACTTCCGCTGTATCCCGAGGGGACCAAC 1099  
Db 1283 AACACTGCCCAGAGGCTGTAGGAAAGGAACAACTTCACTTCACTTCCGAGGGGACCAAT 1342  
QY 1100 GGCAATTGAGGCGCATGTGATGTTCTGGAGGAAATGTTGGCTTCTGGAGAGATGGAC 1159  
Db 1343 GGCACTGAGGAGGGATGTCCGTCACTTGGGACAAAGGCTGTGGTCACTTGGGAAAGATGGAT 1402

QY 1160 GAGAAATGATGTTCTCGCGGTGACCAAGTCAAAATGCTGCCAAAATCTTCAATTTTACCCA 1219  
Db 1403 GAGAACCAAGTTTCTGGCTGTGACCAAGCAACCAATGACCAAAAGTCTTCAACCTTTACCCC 1462  
QY 1220 AGGAAGGGCGAGTGTGCTGTGGCTCTGACGCTGACCTGTGTATATGGAACCCCAAGGCC 1279  
Db 1463 CGGAAGGGCGCATGTTCTGTGGGATCCGATGCGGACCTGTGTATCTGGGACCCCGACAGC 1522  
QY 1280 ACCAAGATCATCTCTGCCAAGACCCACAATCTGAAGGTGGAGTACAACATCTTTCGAGGGA 1339  
Db 1523 GTTAAACCAATCTCTGCCAAGACACACAACAGCTCTCTCGAGTACAACATCTTTCGAGGC 1582  
QY 1340 GTGAGTGCGGGAGCGGCTGCGGTGTCATTAAGTCAAGGCGGAGTGGCGCTGGAGGAC 1399  
Db 1583 ATGAGTGCAGCGGCTCCCACTGTGTGTGTCATCAGCCAGGGGAAGATTTGTCTCTGGAGGAC 1642  
QY 1400 GGAAGATGTTTGTCAACCCCGGGCGGCGCTTCTGCTCGGAAACATTTCCCGGAC 1459  
Db 1643 GGCACCTGTGATGTCACCGAAGGCTCTGAGCGCTACATTTCCCGGAAGCCCTTCCCTGAT 1702  
QY 1460 TTTGTCTACAAGAGGATCAAGCTCGCAACAGGCTGGCGGAGATCCAAGGTGTGCCCGCT 1519  
Db 1703 TTTGTTCACAGGATCAAGGCAAGGACGAGCTGGCTGAGCTGAGAGGGTTCCTCGT 1762  
QY 1520 GGCTGTATGACGGGCGGCTGCAAGAGGTGATGCTGCTGCCAAGCAGGAGTGGCGCT 1579  
Db 1763 GGCCTGTATGACGAGCTGTGTGAAAGTGTCTGTGACGCCCAAGACAGTCACTCCAGCC 1822  
QY 1580 CCGGCGCGGCTCTGCGCCAGGCAAGATCTCCGTGCTCTGCTGGCAACCTACATCAG 1639  
Db 1823 TCCTCGGCCAAGAGCTCTCTGCAAGCAGAGCGGCCCACTGTCCGGAACCTGCACAG 1882  
QY 1640 TCGGGGTTCAGCCCTATCTGGGTCTCAGGCTGATGACACATCGCCCGACGC 1690  
Db 1883 TCTGGATTCAATTTGTCTGTGTCTGATGATGACACAATTTCCCGCGCGC 1933

## RESULT 13

US-10-159-563-8  
; Sequence 8, Application US/10159563  
; Publication No. US2004009154A1  
; GENERAL INFORMATION:  
; APPLICANT: Khan, Javed  
; APPLICANT: Ringner, Markus  
; APPLICANT: Peterson, Carsten  
; APPLICANT: Meltzer, Paul  
; TITLE OF INVENTION: SELECTIONS OF GENES AND METHODS OF USING THE SAME FOR  
; FILE REFERENCE: 11613.56US11  
; CURRENT APPLICATION NUMBER: US/10/159,563  
; PRIOR FILING DATE: 2002-12-09  
; PRIOR APPLICATION NUMBER: US 10/133,937  
; NUMBER OF SEQ ID NOS: 444  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 8  
; LENGTH: 4459  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-159-563-8

Query Match 57.0%; Score 963.8; DB 17; Length 4459;  
Best Local Similarity 73.5%; Pred. No. 1.1e-261;  
Matches 1229; Conservative 0; Mismatches 442; Indels 0; Gaps 0;

QY 20 CCAGAGCAGGAGTGTCTTCCAGGGCAAGAAAAGCATCCCGGATTCAGAGTACCGC 79  
Db 263 CCAGAGCAGAGATGTCTTATCAGGGGAAGAAAATATTCACGCTATCAGAGCATCGT 322  
QY 80 CTTCTGATCAGAGGTGGGAGGATCGTGAATGACGACAGTCTTTTACGCTGTATGCAC 139  
Db 323 CTTCTGATCAAGAGGAGTAAATTTGTTAATGATGACAGTCTGTGTATGACAGATATAC 382

140 GTGGAAGATGGCTGTATATAAACAATCGAGAAAACTCATCGTCCCTGGGGCATCAAG 199  
Db |||||  
383 ATGGAAGATGGCTGTATATAAACAATCGAGAAAACTCATCGTCCCTGGGGCATCAAG 442  
Qy |||||  
200 ACCATTTGAGCGCCACGCGCTGTATGGTCTTCTTCTGGTGGGCTTGAAGCTTCAACAAGGGCTG 259  
Db |||||  
443 ACCATCGAGCGCCACTCCCGATGGTGTATCCCGGAGGAATTGACGTCCACACTCGCTTTC 502  
Qy |||||  
260 CAGATCGCTGTCTGGGCTGACACCGGCTGACGATCTCTGTCTAGGGCCACCAAGGCAGCG 319  
Db |||||  
503 CAGATCGCTGTATCAGGGAATGACGTCTGTGTATGATTTCTTCAGAGAAACCAAGCGCGCC 562  
Qy |||||  
320 CTAGCAGGAGGAACCACTATGATCTTTGGACACGCTCTTCCCGACACGCGGTGTGAGCCTG 379  
Db |||||  
563 CTGGCTGGGGGAACCACTATGATCTTACCACTGTTTCTGTAGCTGGGACAAGCCTG 622  
Qy |||||  
380 CTGGCGGCTTACGAGCAGTGGCGGAGCGGCGGACAGCGCGCTCTGTCTGCACTACTCC 439  
Db |||||  
623 CTGCTGCTTTCACCACTGGAGGAATGGGCGGACAGCAAGTCTCTGTGTGACTACTCT 682  
Qy |||||  
440 CTGCACTGTGACATCAACCGATGGCATGAGGAGCATCAAGGAGAGCTGGAGGCGCTGGTC 499  
Db |||||  
683 CTGCACTGTGACATCAAGCAGTGGCATTAAGGCGCATCCAGGAGAGATGGAAGCGCTGTG 742  
Qy |||||  
500 AAGGAGAGGGGTGTGAATCTCTTCTGTCTTCAATGGCATACAAGGACCGGTGCCAGTGC 559  
Db |||||  
743 AAGGATCAAGGGTAAATTCCTTCTCTGTGTATGCTTTCAGAGTTCGTTCCAGCTA 802  
Qy |||||  
560 AGCGACAGCGAGATGTACGAGATCTTCAAGCATCATCCGGGACTCTGGGGCGCTTGGCCGAG 619  
Db |||||  
803 ACGGATTCAGAGTTATGAAGTACTGAGTGTGATCCGGGATATGGCGCCATAGGCCAA 862  
Qy |||||  
620 GTGCACTGTGAGAACGGGAGCATCTGTGGAGGAGCAGAGCGGTGTCTGGAGCTGGC 679  
Db |||||  
863 GTCACGCGAGAAATTCGCGACATCATTTGACAGAGAGCAGAGGATCTTGGATCTGGCG 922  
Qy |||||  
923 ATCACTGGGCGGCGGAGGAGCATGTCTGAGCGGAGCTGTGAGGAGTCCAGGCGGAGCGTG 982  
Db |||||  
740 TACCGAGCTGTACACATTCGCAAGGAGGCAAACTGCGCGCTGTACGTCACCAAGGATG 999  
Qy |||||  
983 AATCGTGCCATCACTACCAACCAAGAGCACTGCGCGCTGTATATCAAGAGTGATG 1042  
Db |||||  
800 AGCAAGGGGCGGCGGAGCGCATCTGTCAGGCCAAGCGGAGAGGGGTGTGTGTTGG 859  
Qy |||||  
1043 AGCAAGAGCTCTGTGAGGTCTATCGCCGAGGAGCGGAGGAGGAACTGTGTGTATGGC 1102  
Db |||||  
860 GAGCCCATCAACCGCAGCTGGGCGACCGAGGTTTCACTACTGAGAGCAAGAACTGGGCC 919  
Qy |||||  
1103 GAGCCCATCACTGCGAGCTTGGGAAAGGAGCGGCTCCATTTACTTGGAGCAAGAACTGGGCC 1162  
Db |||||  
920 AAGGCTGCACTCTGTGTCAGCGGGGACCTTCAGGTGACAGGCGGCGGCACTGCACTTC 1039  
Qy |||||  
1163 AAGGCTGCTGCTTGTCTACCTCCCGACCTTGGAGGCTGATCCAGCACTCCAGACTTT 1222  
Db |||||  
980 CTCACTGTCTGTCTGTCAGCGGGGACCTTCAGGTGACAGGCGGCGGCACTGCACTTC 1039  
Qy |||||  
1223 CTCACTGTCTGTCTGTCAGCGGGGACCTTCAGGTGACAGGCGGCGGCACTGCACTTC 1282  
Db |||||  
1040 ACCACTGCGCAGAGGCTGTGGGCAAGGACAACTTGGCGTGTATCCCGAGGCGGCAAGC 1099  
Qy |||||  
1283 AACACTGCGCAGAGGCTGTGGGCAAGGACAACTTGGCGTGTATCCCGAGGCGGCAAGC 1342  
Db |||||  
1100 GGCAATTGAGAGCGCATGTCTGATGTGTGGGAAATGTGTGGCTCTTGGGAAAGATGGAC 1159  
Qy |||||  
1343 GGCACTGAGAGCGGATGTCCGTCTATCTGGGCAAGGCTGTGTCTGCTGAGGAGATGGAT 1402  
Db |||||  
1160 GAGCAATGATTTCTGCGGCTGACAGTACAAATGCTGCGCAAAATCTTCAATTTTACCCA 1219  
Qy |||||  
1403 GAGAAACGATTTGTGGCTGTGACAGCAACCAATGCGAGGCAAGTCTTCAACCTTTACCCC 1462  
Db |||||  
1220 AGGAAGGGGCGAGTGGCTGTGGGCTCTGACGCTGACCTGTGTATATGGAACCCCAAGGCC 1279  
Qy |||||

1463 CGGAAGGCGCATGCTGTGTGGATCCGATCCGACCTGCTCATCTGGGACCCCGACGC 1522  
Db |||||  
1280 ACCAAGATCATCTCTGCGCAAGACCCACATCTGAAGCTGGAGTACAAATCTTCGAGGA 1339  
Qy |||||  
1523 GTTAAACCATCTCTGCGCAAGACACACACAGCTCTCTCGAGTACAAATCTTTGAAGGC 1582  
Db |||||  
1340 GTGGAGTGCAGGAGCGGCTGCGGTGTGTCATAAGTTCAGGGCGGAGTGGCGCTGGAGGAC 1399  
Qy |||||  
1583 ATGGAGTGCAGGCGGCTCCCACTGTGTGTGTCATGAGCAGGGGAAGATTTCTCTGGAGGAC 1642  
Db |||||  
1400 GGAAGATGTTTGTTCACCCCGGGCGGCGCTTCTGTCCTCGGAAAAACATTTCCCGGAC 1459  
Qy |||||  
1643 GGCACCTTCATGTTCACCGAAGGCTCTGGAAGCTTATTCCTCGGAGGCGCTTCTCTGAT 1702  
Db |||||  
1460 TTTGTCTCAAGAGAGTCAAAAGCTCGCAACAGGCTGGCGGAGATCCACGGTGTGCCCGT 1519  
Qy |||||  
1703 TTTGTCTCAAGAGTCAAAAGCTCGCAACAGGCTGGCGTGGCTGAGCTGAGAGGGGTTCCTGT 1762  
Db |||||  
1520 GGGCTGTATGACGGGCGGCTCCACGAGGTGATGCTGCTGCCAAGCCAGGAGTGGCGCT 1579  
Qy |||||  
1763 GGCCTGTATGACGAGGCTCTGCTGCTGAGGAGGCGGCGGCTGCTGAGGAGGCTTCCAGGC 1822  
Db |||||  
1580 CCGGCGCGGCGCTCTGCGCCAGGCAAGATCTCCGTGCTCTCTGTGCGCAACCTTACATCAG 1639  
Qy |||||  
1823 TCTCGGCGCAAGAGGCTCTCTGCGAAGCAGGAGGCGGCGGCTGCTGAGGAGGCTTCCAGGC 1882  
Db |||||  
1640 TGGGGTTCAGGCTATCTGGGCTCAGGCTGATGACCATCGCCCGAGCGC 1690  
Qy |||||  
1883 TCTGGATTTCAGTCTGTCTGTGCTCAGATTGATGACAAATTTCCCGCGCGC 1933  
Db |||||

RESULT 14  
US-10-205-331-53  
; Sequence 53, Application US/10205331  
; Publication No. US20040058326A1  
; GENERAL INFORMATION:  
; APPLICANT: Warner-Lambert Company  
; APPLICANT: Lee, Kevin  
; APPLICANT: Dixon, Alistair  
; APPLICANT: Brooksbank, Robert  
; APPLICANT: Pinnoch, Robert  
; TITLE OF INVENTION: Identification and Use of Molecules Implicated in Pain  
; FILE REFERENCE: WL-A-018199  
; CURRENT APPLICATION NUMBER: US/10/205,331  
; CURRENT FILING DATE: 2002-07-24  
; PRIOR APPLICATION NUMBER: GB 0118354.0  
; PRIOR FILING DATE: 2001-07-27  
; NUMBER OF SEQ ID NOS: 117  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 53  
; LENGTH: 4459  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; OTHER INFORMATION: Dihydropyrimidinase-related protein  
US-10-205-331-53

Query Match 57.0%; Score 963.8; DB 18; Length 4459;  
Best Local Similarity 73.5%; Pred. No. 1.1e-261;  
Matches 1229; Conservative 0; Mismatches 442; Indels 0; Gaps 0;

Qy 20 CCNAGGAGGAGGATGCTCTCCAGGCGAAGAAAGCATCCCGGATCAGGATGACCGC 79  
Db 263 CCAGAGAGAGATGCTTATCAGGGGAAGAAATATTTCCAGCATCAGGCGATCGT 322  
Qy 80 CTTCTCATCAGAGGTGGGAGGATCGTGAATGACGACAGTCTCTTTTACGCTGATGTGAC 139  
Db 323 CTTCTCATCAGAGGAGTAAATTTTATGATGACCACTCGTCTTCTATGACAGATATAC 382  
Qy 140 GTGGAAGATGGCTGTGATATAAACAATCGGAGAAAACTCATCGTCTCCCTGGGGCATCAAG 199  
Db 383 ATGGAAGATGGGTTGATCAAGCAAAATAGGAGAAATCTGATTTGTGCCAGGAGTGAAG 442

```
QY 200 ACCATTGACGCCACGGCCTGATGGTCTTCTGGTGGCGTTGACGTCCACAAAGGCTG 259
Db 443 ACCATCAGAGGCCACTCCCGGATGGTATCCCGGAGGAATTGACGTCCACACTCGTTC 502
QY 260 CAGATGCTGTCTGGGCAATGACACCGGCTGACGACTTCTGTGAGGCAACCAAGGACGG 319
Db 503 CAGATGCTGTATCAGGGAATGACGCTGCTGTGATGATTTCTTCCAAAGGAACAAAGGCGGCC 562
QY 320 CTAGCAGAGGAACCAACCATGATCTTGGACCAACGCTTTCGCCGACACGGGTGTGAGCCTG 379
Db 563 CTGCTGGGGGACCACTATGATCATTTGACCAACGTTGTTCTTGAGCTTGGGCAAGGCTG 622
QY 380 CTGGCGGCTTACGAGCAGTGGGAGACGGGCGGACAGCGCGCTCTGCTGCGACTACTCC 439
Db 623 CTGCTGCTCTTGAACCAAGTGGAGGAATGGGCGGACACAGCAAGTCTCTGTGACTACTCT 682
QY 440 CTGCAGCTGACATCACCCGATGCGATGAGAGCATCAAGGAGGAGCTGAGGCGCTTGTC 499
Db 683 CTGATGTGACATCACGAGTGGCATTAAGGGCATCCAGGAGGAGATGGAAGCGCTTGTG 742
QY 500 AAGGAAGAGGTGTGAATCTCTTCTGCTCTTCAATGCGATACAGGACCGGTGCCAGTGC 559
Db 743 AAGGATCAGGGGTAAATCTCTCTGCTGTGATGATGCTTTCAAAGATCGCTTCCAGTGA 802
QY 560 AGCGACAGCAGATGTACGAGATCTTCAGCATCATCCGGGACCTGGGGCTTTGGGCCAG 619
Db 803 ACGGATTCAGAGATTTATGAAGTACTGAGTGTGATCCGGGATATTTGGCGCCATAGGCCAA 862
QY 620 GTGCAGCTGAGAACGGGACATCGTGGAGGAGGAGCAGAGCGGTTGCTGGAGCTCGGC 679
Db 863 GTCCACGAGAAATGGCGACATCATTTGCGAGAGGAGCAGAGGATTCCTGGATCTGGGC 922
QY 680 ATCACTGGCCCGGAGGCCACGTCTCAGCCACCCCGAGGAGTGGAGGCTGAGGCGGTG 739
Db 923 ATCAGGGCCCGGAGGACATGTGCTGAGCGGACCTGAGGAGTCCAGGGCGGAGCGGTG 982
QY 740 TACCGAGCTGTCAACATCGCAAGCAGGCAAACTGCCGCTGTACGTTCACCAAGGATGATG 799
Db 983 AATCGTGCCATCACATCGCCACACAGCAACCACTGCCCGCTGTATATCACCAAGGTGATG 1042
QY 800 AGCAAGGGGGCGGCGAGGATCGCTCAGGCCAGCGCAGAGGGGTGCTGTTTGGG 859
Db 1043 AGCAAAAGCTCTGTGAGGTCATCGCCACGAGCGACGGAAGGGAACGTGTGTGTATGCGC 1102
QY 860 GAGCCCATCACCGCAGCTGGGACGACGAGCTTTCACACTACTGAGCAGAGAACTGGGCC 919
Db 1103 GAGCCCATCACTGCGAGCTTGGGAACGAGCGGCTCCCATTTACTGGAGCAAGAACTGGGCC 1162
QY 920 AAGGCTGACGCTTGTGTCATCACATCACCCCTGTCAACCCAGACCCCAACCGGACGACAC 979
Db 1163 AAGGCTGCTGCTTGTGTCATCTCCCACTCCCACTTGGAGCTTGTATCCACCACTCCAGACTTT 1222
QY 980 CTCACTGCTGTGTCTCAGGGGACCTCAGGTGACAGGAGCGGCCCACTGACACTTC 1039
Db 1223 CTCACTCTTGTGTCTGTGAGACCTCAGGTCAAGGCTCAGGGGAGTGGCCATTCAGCGTTT 1282
QY 1040 ACCATGCCCAGAGGCTGTGGGCAAGGACAACTTTCGCGCTGATCCCGAGGCGACCAAC 1099
Db 1283 AACACTGCCAGAGGCTGTAGGAAGAGCAACTTTCACCTGATTCGGAGGGGACCAAT 1342
QY 1100 GGCATTGAGAGGCGCATGTGATGCTGTGGAGAAATGTGTGGCTCTCGGGAAGATGGAC 1159
Db 1343 GGCATGAGAGCGGATGTCGTCATCTGGGACAGGCTGTGTCTGTCATCGGAGATGGAT 1402
QY 1160 GAGATGAGTGTGTGGGCTGACCAAGTAAATGCTGCCAAAATCTTCAATTTTACCCA 1219
Db 1403 GAGAACCAAGTGTGTGGCTGTGACCAAGCAACCAATGACGCAAAAGTCTTCAACCTTTACCCC 1462
QY 1220 AGGAGGGGAGTGGCTGTGGCTGTGACCTGAGCTGATATGGAACCCCAAGGCC 1279
Db 1463 CGGAAAGGCCCATGTGCTGGGATCCGATGCGGACCTGCTGATCTGGGACCCCGACAGC 1522
```

```
QY 1280 ACCAAGATCATCTCTGCCAAGACCCACAAATCTGAAAGTGGAGTACAAATCTTTCAGGGA 1339
Db 1523 GTTTAAACCATCTCTGCCAAGACACAAACAGCTCTCTCGAGTACAAATCTTTGAAGGC 1582
QY 1340 GTGGAGTCCGGGAGCGCTGCTGATCAAGTCAAGGCGAGTGGCGCTGGAGGAC 1399
Db 1583 ATGGAGTCCGGGCTCCCACTGGTGGTCAACGCGAGGGGAAGATTGTCTCTGGAGGAC 1642
QY 1400 GGAAGAATGTTTCTACCCCGGGGGCGGCTTCTGCTCCCTCGGAAAAATTTCCCGGAC 1459
Db 1643 GGCACCTGCTATCTACCGAAGGCTCTGGAAGCTACATTTCCCGGAAGCCCTTCCCTGAT 1702
QY 1460 TTTGTCTACAAAGAGTCAAGCTCGCAACAGGCTGGGAGATCCACGCTGTGTCGCCGT 1519
Db 1703 TTTGTTTACAAAGGTATCAAGGCAAGGAGCAGGCTGGCTGAGCTGAGAGGGGTCTCTCGT 1762
QY 1520 GGGCTGTATGAGCGGCGCTGCCACAGGCTGATGCTGCTGCCAAGCCAGGAGTGGCGCT 1579
Db 1763 GGCCTGTATGAGCGAAGCTGTGTGAGTGTCTGTGAGCCCAAGACAGTCACTTCCAGCC 1822
QY 1580 CCGGCCCGCGCTCTGCCAGGCAAGATCTCGTGTGCTCTGTGGCGCAACCTTACATCAG 1639
Db 1823 TCCTCGGCCAAGAGGCTCTCTGCCAAGCAGCAGGCCCCACCTGTCCGGAACCTGCACAG 1882
QY 1640 TCGGGGTTCAAGCTTATCTGAGTCTCAGGCTGTATGACATTCGCCCGACGC 1690
Db 1883 TCTGATTCAGTTTGTCTGCTGCTCAGATTGATGACAAATTTCCCGCGCGC 1933
```

## RESULT 15

```
US-10-788-792-1
; Sequence 1, Application US/10788792
; Publication No. US20040191819A1
; GENERAL INFORMATION:
; APPLICANT: Bayer Pharmaceuticals Corporation
; APPLICANT: Eveleigh, Deepa
; APPLICANT: Bigwood, Douglas
; TITLE OF INVENTION: EXPRESSION PROFILES FOR BREAST CANCER AND METHODS OF USE
; FILE REFERENCE: 5152
; CURRENT APPLICATION NUMBER: US/10/788,792
; CURRENT FILING DATE: 2004-02-27
; PRIOR APPLICATION NUMBER: US 60/450,655
; PRIOR FILING DATE: 2003-02-28
; NUMBER OF SEQ ID NOS: 254
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 1
; LENGTH: 4459
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-788-792-1
```

```
Query Match 57.0%; Score 963.8; DB 19; Length 4459;
Best Local Similarity 73.5%; Pred. No. 1.1e-261;
Matches 1229; Conservative 0; Mismatches 442; Indels 0; Gaps 0;
```

```
QY 20 CCAGAGCAGAGGATGTCTTCCAGGGCAAGAAAGCATCCCGGATCAGAGTGACCGC 79
Db 263 CCAGAGGAGAGATGTCTTATCAGGGGAGAGAAATATTCACGCTATCAGAGGATCGT 322
QY 80 CTTCTGATCAGAGGTGGGAGGATCGTGAATGACGACAGCTCTTTTACGCTGATGTGCAC 139
Db 323 CTTCTGATCAAAGGAGGTAAATTTTAATGATGACGACAGCTGCTTATGACGACATATAC 382
QY 140 GTGGAAGATGGCTTGTAAATAAACAATCGGAGAAACCTCATCTCTCTGGGCGATCAAG 199
Db 383 ATGGAAGATGGGTGTATCAAGCAATAGGAGAAATCTGATTGTCCAGGAGGAGTGAAG 442
QY 200 ACCATTGAGCGCCCAACCGCTGTGCTCTCTGCTGGGCTTTCAGCTCCACACAAGGCTG 259
Db 443 ACCATCGAGGCCCACTCCCGGATGGTGTATCCCGGAGGAAATGACGCTCCACACTCGTTTC 502
QY 260 CAGATGCTGTCTCTGGGCAATGACACCGGCTGACGACTTCTGTGAGGCGACCAAGGACGG 319
```

Db 503 CAGATGCTGATCAGGAAATGACGTCTGCTGATGATTTCTTCCAGGAACCAAGGCGGC 562  
Qy 320 CTAGCAGGAGAACCACTGATCTTGGACCAAGTCTTCCCGACACAGGGTGTGACCTTG 379  
Db 563 CTGGCTGGGGGAACCACTATGATCATTTGACCAAGTGTCTTCTGAGCTGGGAACAAGCCTG 622  
Qy 380 CTGGCGGCTTACGACGATGGCGGAGCGGGGACAGCGCGGCTGCTGCGACTACTCC 439  
Db 623 CTGCTGCTTTCAGCAGTGGAGGGAATGGGCCGACAGCAAGTCTCTGCTGTGACTACTCT 682  
Qy 440 CTGCACTGGACATCAACCGATGGCATGAGAGCATCAAGAGAGAGCTGGAGGCCCTGGTC 499  
Db 683 CTGCATGTGGACATCAGCGAGTGGCATAGGGCATCAGGAGGAGATGGAAGCGCTTGTG 742  
Qy 500 AAGGAGAGGGTGTGAACCTCTCTGCTGCTTCACTGGCATACAAGACCGGTGCGAGTGC 559  
Db 743 AAGGATCAGCGGGTAAATTTCTTCTCGTGTACATGGCTTTCAAAGATCGCTTCCAGCTA 802  
Qy 560 AGCGACAGCCAGATGTACGAGATCTTCAGCATCATCCGGACCTGGGGGCTTTGGGCCAG 619  
Db 803 ACGGATTGCCAGATTTATGAGTACTGAGTGTGATTCGGGATATTTGGGCCATAGGCCAA 862  
Qy 620 GTGCACTGTAGAACCGGGACATCTGTGGAGGAGAGCAGAACGGTGTCTGTGGAGCTCGGC 679  
Db 863 GTCCACGCAGAAAATGGCGACATCATTTGCAGAGGAGCAGCAGAGGATCCTTGGATCTGGGC 922  
Qy 680 ATCACTGGCCCCGAGGCCACGTGCTCAGCCACCCCGAGGAGGTGAGGCTGAGGCGGTG 939  
Db 923 ATCACGGGCCCCGAGGGACATGTGCTGAGCCGACCTGAGGAGGTGAGGCCGGAAGCCGTG 982  
Qy 740 TACCGAGCTGTACACATCGCAAGCAGGCAAACTGCGCGCTGTACGTACCAAGAGTGTATG 799  
Db 983 AATCGTGCCATCACTATCGCAACCAAGACCACTGCGCGCTGTATATCACAAGGATGATG 1042  
Qy 800 AGCAAGGGGGCGCCGACGCGCATCTGCTCAGGCCAAGCGCAGAGGGGTGTGCTTTTGGG 859  
Db 1043 AGCAAAAGCTCTGCTGAGGTCTATCGGCCAGGCACGGAAGGGAACCTGTGGTGTATGGC 1102  
Qy 860 GAGCCCATCACCGCCAGCTTGGGCAACGAGGTTTCACTACTGGAGCAAGAACTGGGCC 919  
Db 1103 GAGCCCATCATCTGCCAGCTTGGGAAACGAGCGGCTCCCATTTACTGGAGCAAGAACTGGGCC 1162  
Qy 920 AAGGCTGCAGCTTCTGTACATCAACCCCTGTCAACCCAGACCCACACCGCAGACCCAC 979  
Db 1163 AAGGCTGTGCTTGTCTACTCTCCACCCCTTGAGCGCTGTATCAACCACTCCAGACTTT 1222  
Qy 980 CTCACTGTCTGTCTCAGCGGGGACCTCCAGGTGACAGGCGGCCACTGCACTTC 1039  
Db 1223 CTCAACTCTCTGTCTCTGTGGAGACCTCCAGGTCAAGGGCAGTGGCCATTCACGTTT 1282  
Qy 1040 ACCACTGCCCAGAGGCTGTGGGCAAGGACAACTTCCGCTGTATCCCGAGGGGACCCAC 1099  
Db 1283 AACACTGCCCAGAGGCTGTAGAAAGGACAACTTCACTCCCTGATTCGGAGGGGACCAAT 1342  
Qy 1100 GGCATTTGAGGAGCGCATGTGATGTCTGGGAGAAATGTGTGGCTCTTGGGAAAGATGGAC 1159  
Db 1343 GGCATCTGAGAGCGGATGTCTCGTCACTTGGGACAGGCTGTGTCTACTGGGAAAGATGGAT 1402  
Qy 1160 GAGAATGAGTTCTGCGGGTGACAGTACAAATGCTGCCAAAATCTTCAAATTTTACCCA 1219  
Db 1403 GAGAACCAAGTTTGTGGCTGTGACCAAGCACTGACCAAAAGTCTTCAACCTTTTACCCC 1462  
Qy 1220 AGGAAGGGGCGAGTGGCTGTGGGCTGTGACGCTGACCTGTGATATGGAAACCCCAAGGCC 1279  
Db 1463 CGAAAAGGCCGCAATGCTGTGGGATCCGATGCGGACCTGTGCTATCTGGGACCCCGACAGC 1522  
Qy 1280 ACCAAGATCATCTGTGCCAAGACCAAACTGAAAGTGGAGTACAACTCTTTCGAGGGA 1339  
Db 1523 GTTAAACCACTCTGTGCCAAGACACACACAGCTCTCTCGAGTACAACTCTTTGAGGC 1582  
Qy 1340 GTGAGTGGCGGGAGCGCTGCGGTGTCATTAAGTCAGGGCCGAGTGGCGCTGGAGGAC 1399  
Db 1583 ATGGAGTGGCGGCTCCCCACTGGTGGTTCATCAGCCAGGGGAAGATTGTCTCTGGAGGAC 1642

Qy 1400 GGGAAGATGCTTTGTCAACCCCGGGGCGGCGCTTCTGTCCTCGGAAAAACATTTCCGGAC 1459  
Db 1643 GGAACCTGCATGTCAACCGAAGGCTCTGACACGTTACATTTCCCGGAAGCCCTTCCCTGAT 1702  
Qy 1460 TTTGTCTACAAGAGGATCAAAAGCTCGCAACAGGCTGGCGGAGATCCACGTTGTGCCCCGT 1519  
Db 1703 TTTGTTTACAAGGATCAAGGCAAGGAGCAGGCTGGCTGAGCTGAGAGGGGTTCCTCGT 1762  
Qy 1520 GGGCTGTATGACGGGCGGCTCCACGAGGTGATGGTGTCTGCGCAAGCCAGGGAGTGGCGCT 1579  
Db 1763 GGCCTGTATGACGGACCTGTGTGTGAAGTGTCTGTGACGCCCAAGACAGTCACTCCAGCC 1822  
Qy 1580 CCGGCCCGCGCTTCTTGGCCAGCAAGATCTCGGTGCTCTGTGCGCAACCTACATCAG 1639  
Db 1823 TCTCGGCCCAAGACGCTCTCTGCCAAGCAGGAGGCCACCTGTCCGGAACTTGCACCAAG 1882  
Qy 1640 TCGGGGTTTCAGGCTATCTGGGTCTCAGGCTGATGACCAATCGCCCGACGC 1690  
Db 1883 TCTGGATTTCAGTTTGTCTGTGCTCAGATTGATGACAAATTTCCCGCGCGC 1933

Search completed: September 24, 2005, 16:52:25  
Job time : 1087 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 24, 2005, 14:19:36 ; Search time 1390 Seconds  
(without alignments)  
210.726 Million cell updates/sec

Title: US-09-367-496C-8

Perfect score: 2982

Sequence: 1 MSFQKKSIPRITSDLLR.....RTAQKIMAPPGRSNTISLS 572

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Uniprot\_03.\*

1: uniprot\_sprot.\*

2: uniprot\_trembl.\*

Pred. No. is the number of results predicted by chance to have a  
' score greater than or equal to the score of the result being printed,  
' and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2978	99.9	572	1 DPY4 HUMAN	O14531 homo sapien
2	2824	94.7	572	1 DPY4 MOUSE	O35098 mus musculu
3	2795	93.7	564	1 DPY4 RAT	O62951 rattus norv
4	2519	84.5	574	2 Q9D5Z6	O9dd26 gallus gall
5	2488	83.4	650	2 Q71SG0	Q71sg0 gallus gall
6	2441.5	81.9	668	2 Q7ZTP7	Q7ztp7 xenopus lae
7	2368	79.4	572	1 DPY2 CHICK	Q90635 gallus gall
8	2361	79.2	572	2 Q71SG1	Q71sg1 gallus gall
9	2353	78.9	572	1 DPY2 BOVIN	O02675 bos taurus
10	2349	78.8	572	1 DPY2 RAT	P47942 rattus norv
11	2349	78.8	572	2 Q6P5D0	Q6p5d0 mus musculu
12	2345	78.6	572	1 DPY2 HUMAN	Q16555 homo sapien
13	2331	78.2	572	1 DPY2 MOUSE	O08553 mus musculu
14	2315	77.6	676	2 Q71SG2	Q71sg2 gallus gall
15	2276	76.3	573	2 Q6PB02	O6pb02 xenopus lae
16	2215.5	74.3	571	2 Q6GL72	O6gl72 xenopus tro
17	2209.5	74.1	571	2 Q640K6	Q640k6 xenopus lae
18	2203.5	73.9	571	2 Q7ZQW8	Q7zqw8 xenopus lae
19	2198	73.7	570	2 Q71SF8	Q71sf8 gallus gall
20	2193	73.5	570	1 DPY3 MOUSE	O62188 mus musculu
21	2192	73.5	570	1 DPY3 HUMAN	O14195 homo sapien
22	2191	73.5	570	2 Q81XW6	Q81xw6 homo sapien
23	2191	73.5	570	2 Q91XM8	Q91xm8 rattus norv
24	2178.5	73.1	571	1 DPY3 XENLA	O13022 xenopus lae
25	2174	72.9	572	2 Q71SG3	Q71sg3 gallus gall
26	2163	72.5	572	1 DPY1 MOUSE	P97427 mus musculu
27	2158	72.4	572	1 DPY1 RAT	O62950 rattus norv
28	2157	72.3	572	1 DPY1 HUMAN	Q14194 homo sapien
29	2138	71.7	683	2 Q8K4H3	Q8k4h3 rattus norv
30	2135	71.6	684	2 Q6DEN2	O6den2 homo sapien
31	2117	71.0	683	2 Q71SG4	Q71sg4 gallus gall

## RESULT 1

ID DPY4 HUMAN STANDARD; PRT; 572 AA.

AC O14531; O00240;

DT 15-JUL-1998 (Rel. 36, Created)

DT 15-JUL-1998 (Rel. 36, Last sequence update)

DE Dihydropyrimidinase-related protein-4 (DRP-4)

DE mediator protein 3) (CRMP-3) (UNC33-like phosphoprotein 4) (ULIP4

DE protein).

GN Names=DPYSL4; Synonyms=CRMP3, ULIP4;

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OX NCBI\_TaxID=9606;

RN [1]

RC TISSUE=Brain;

RA Hamajima N., Kato Y., Kouwaki M., Wada Y., Sasaaki M., Nonaka M.;

RL Submitted (AUG-1997) to the EMBL/GenBank/DBJ databases.

RN [2]

RP SEQUENCE OF 1-553 FROM N.A.

RC TISSUE=Retina;

RX MEDLINE=98314496; PubMed=9652388;

RA Byk T., Ozon S., Sobel A.;

RT "The Ulip family phosphoproteins -- common and specific properties.";

RL Eur. J. Biochem. 254:14-24(1998).

RN [3]

RP RECOGNITION BY ANTI-CV2 AUTOANTIBODIES IN PND.

RC PubMed=10594648;

RA Honnorat J., Byk T., Kusters I., Aguera M., Ricard D., Rogemond V.,

Quach T.T., Aunis D., Sobel A., Mattei M.-G., Kolattukudy P.E.,

Belin M.-P., Antoine J.-C.;

RT "Ulip/CRMP proteins are recognized by autoantibodies in paraneoplastic

neurological syndromes.";

RL Eur. J. Neurosci. 11:4226-4232(1999).

CC -I- FUNCTION: May be involved in the axonal growth during development

CC by transducing signals from different semaphorins.

CC -I- SUBCELLULAR LOCATION: Cytoplasmic (Potential).

CC also called anti-CV2 autoantibodies, are present in sera from

patients with paraneoplastic neurological diseases (PND). PND are

disorders of the nervous system associated with various systemic

cancers which are not a direct result of the tumor mass or

metastasis, but attributed to remote effects of the cancer.

CC -I- SIMILARITY: Belongs to the dehydropyrimidinase family.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

between the Swiss Institute of Bioinformatics and the EMBL outstation -

the European Bioinformatics Institute. There are no restrictions on its

use by non-profit institutions as long as its content is in no way

modified and this statement is not removed. Usage by and for commercial

entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

Q6pij1 mus musculu  
Q86u75 homo sapien  
Q14117 homo sapien  
Q63150 rattus norv  
Q642f0 rattus norv  
Q9eqf5 mus musculu  
Q961l1 homo sapien  
Q9eqf6 mus musculu  
Q9jhu0 rattus norv  
Q9bpue homo sapien  
Q62952 rattus norv  
Q9jng8 rattus norv  
Q86lt3 drosophila  
Q7qbkg9 anopheles g

## ALIGNMENTS

32 2110 70.8 686 2 Q6PIJ1  
33 1691 56.7 619 2 Q86U75  
34 1556 52.2 519 1 DPVS HUMAN  
35 1550 52.0 519 1 DPVS\_RAT  
36 1550 52.0 519 2 Q642F0  
37 1520 51.0 519 1 DPVS\_MOUSE  
38 1492 50.0 404 2 Q961L1  
39 1481.5 49.7 564 1 DPVS\_MOUSE  
40 1481.5 49.7 564 1 DPVS\_RAT  
41 1469.5 49.3 564 1 DPVS\_HUMAN  
42 1457 48.9 358 1 DPVS\_RAT  
43 1456 48.8 563 2 Q9JNG8  
44 1347 45.2 594 2 Q86LT3  
45 1318 44.2 595 2 Q7QBK9

```

CC -----
DR EMBL; AB006713; BAA21886.1; -.
DR EMBL; Y10976; CAA71872.1; -.
DR HSSP; Q45515; 1K1D.
DR MEROPS; M38.977; -.
DR Genew; HGNC:3016; DPYSL4.
DR MTM; 608407; -.
DR GO; GO:0007399; P. neurogenesis; TAS.
DR InterPro; IPR006680; Amidohydro_1.
DR InterPro; IPR005847; DHOase.
DR Pfam; PF01979; Amidohydro_1; 1.
DR ProDom; PD000518; DHOase; 1.
FT CONFLICT 122 122 (in Ref. 2).
SQ SEQUENCE 572 AA; 61905 MW; 3E7A33E3BBD5B8E9 CRC64;

Query Match 99.9%; Score 2978; DB 1; Length 572;
Best Local Similarity 99.8%; Pred. No. 5.1e-207;
Matches 571; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSFQKKSIPIRISDRLLGRGRVNDQSFYADVHVEDGLIKQIGENLIVPGGIKTIDA 60
DB 1 MSFQKKSIPIRISDRLLGRGRVNDQSFYADVHVEDGLIKQIGENLIVPGGIKTIDA 60
QY 61 HGLMVLPGGVVHTRIQMPVLGMPVLTADDPCQGTKAALAGGTTMILDHVPDGTGVSLLAY 120
DB 61 HGLMVLPGGVVHTRIQMPVLGMPVLTADDPCQGTKAALAGGTTMILDHVPDGTGVSLLAY 120
QY 121 EQRERADSAACDYSILHVDITRWHSIKKELEALVKEGVNSFLVPMAYKDRCCQSDSQ 180
DB 121 EQRERADSAACDYSILHVDITRWHSIKKELEALVKEGVNSFLVPMAYKDRCCQSDSQ 180
QY 181 MYEIPSIIRDLGALAOVHAENGDI VEEBQKRLLELGTGPBGHVLSPHEVEAEVYRAV 240
DB 181 MYEIPSIIRDLGALAOVHAENGDI VEEBQKRLLELGTGPBGHVLSPHEVEAEVYRAV 240
QY 241 TIQAQNCPLVYTKVMSKGAADVAIAQAKRGVVVFGEPITASLGTGSHYWSKNWAKAA 300
DB 241 TIQAQNCPLVYTKVMSKGAADVAIAQAKRGVVVFGEPITASLGTGSHYWSKNWAKAA 300
QY 301 FVTSPPVNPDPPTADHLTCLSSGDLQVTSAGHCTFTTAQKAVGKDNFALIPGTGIER 360
DB 301 FVTSPPVNPDPPTADHLTCLSSGDLQVTSAGHCTFTTAQKAVGKDNFALIPGTGIER 360
QY 361 RMSMWKCVASGKMDNEFVAVTSTNAAKIFNFPYPRKGRVAVGSDADLVIMNPKATKII 420
DB 361 RMSMWKCVASGKMDNEFVAVTSTNAAKIFNFPYPRKGRVAVGSDADLVIMNPKATKII 420

QY 421 SAKTHNLNVEYNIPEGVECHGAPAVVISQGRVALEDGKMFVTPGAGRFVPRKTFPPDVPYK 480
DB 421 SAKTHNLNVEYNIPEGVECHGAPAVVISQGRVALEDGKMFVTPGAGRFVPRKTFPPDVPYK 480
QY 481 RIKARNRLAEIHGVPRLGYPGVHVMVPAKPGSGAPARASCPOKISVPPVRNLHSGGFS 540
DB 481 RIKARNRLAEIHGVPRLGYPGVHVMVPAKPGSGAPARASCPOKISVPPVRNLHSGGFS 540
QY 541 LSGSQADHDHIAARTAKIMAPPGSRNITSLS 572
DB 541 LSGSQADHDHIAARTAKIMAPPGSRNITSLS 572

RESULT 2
ID DPY4 MOUSE STANDARD; PRT; 572 AA.
AC Q35098; O08886;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DE 03-JUL-2004 (Rel. 44, Last annotation update)
DE Dihydropyrimidinase related protein-4 (DRP-4) (Collapsin response
DE mediator protein 3) (CRMP-3) (UNC33-like phosphoprotein 4) (ULIP4
DE protein).
GN Names=Dpy414; Synonyms=Crm3, Ulp4;
OS Mus musculus (Mouse).

```

```

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Embryo;
RA Hamajima N., Kato Y., Kowaki M., Wada Y., Sasaki M., Nonaka M.;
RL Submitted (AUG-1997) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=ICR; TISSUE=Brain;
RX MEDLINE=98314496; PubMed=9652388;
RA Byk T., Ozon S., Sobel A.;
RT "The Ulip family phosphoproteins -- common and specific properties.";
RL Eur. J. Biochem. 254:14-24(1998).
CC -!- FUNCTION: May be involved in the axonal growth during development
CC by transducing signals from different semaphorins.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (Potential).
CC -!- SIMILARITY: Belongs to the dehydroxyrimidine family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AB006715; BAA21888.1; -.
DR EMBL; Y09079; CAA70299.1; -.
DR HSSP; Q45515; 1K1D.
DR MEROPS; M38.977; -.
DR MGD; MGI:1349764; Dpysl4.
DR InterPro; IPR006680; Amidohydro_1.
DR InterPro; IPR005847; DHOase.
DR InterPro; IPR011059; Metallo_hydrolase.
DR Pfam; PF01979; Amidohydro_1; 1.
DR ProDom; PD000518; DHOase; 1.
FT CONFLICT 125 126 ER -> DG (in Ref. 2).
FT CONFLICT 354 354 G -> V (in Ref. 2).
FT CONFLICT 420 420 F -> I (in Ref. 2).
SQ SEQUENCE 572 AA; 61961 MW; 37671129FC02C7AF CRC64;

Query Match 94.7%; Score 2824; DB 1; Length 572;
Best Local Similarity 93.2%; Pred. No. 7.3e-196;
Matches 533; Conservative 21; Mismatches 18; Indels 0; Gaps 0;

QY 1 MSFQKKSIPIRISDRLLGRGRVNDQSFYADVHVEDGLIKQIGENLIVPGGIKTIDA 60
DB 1 MSFQKKSIPIRISDRLLGRGRVNDQSFYADVHVEDGLIKQIGENLIVPGGIKTIDA 60
QY 61 HGLMVLPGGVVHTRIQMPVLGMPVLTADDPCQGTKAALAGGTTMILDHVPDGTGVSLLAY 120
DB 61 HGLMVLPGGVVHTRIQMPVLGMPVLTADDPCQGTKAALAGGTTMILDHVPDGTGVSLLAY 120
QY 121 EQRERADSAACDYSILHVDITRWHSIKKELEALVKEGVNSFLVPMAYKDRCCQSDSQ 180
DB 121 EQRERADSAACDYSILHVDITRWHSIKKELEALVKEGVNSFLVPMAYKDRCCQSDSQ 180
QY 181 MYEIPSIIRDLGALAOVHAENGDI VEEBQKRLLELGTGPBGHVLSPHEVEAEVYRAV 240
DB 181 MYEIPSIIRDLGALAOVHAENGDI VEEBQKRLLELGTGPBGHVLSPHEVEAEVYRAV 240
QY 241 TIQAQNCPLVYTKVMSKGAADVAIAQAKRGVVVFGEPITASLGTGSHYWSKNWAKAA 300
DB 241 TIQAQNCPLVYTKVMSKGAADVAIAQAKRGVVVFGEPITASLGTGSHYWSKNWAKAA 300
QY 301 FVTSPPVNPDPPTADHLTCLSSGDLQVTSAGHCTFTTAQKAVGKDNFALIPGTGIER 360
DB 301 FVTSPPVNPDPPTADHLTCLSSGDLQVTSAGHCTFTTAQKAVGKDNFALIPGTGIER 360
QY 361 RMSMWKCVASGKMDNEFVAVTSTNAAKIFNFPYPRKGRVAVGSDADLVIMNPKATKII 420
DB 361 RMSMWKCVASGKMDNEFVAVTSTNAAKIFNFPYPRKGRVAVGSDADLVIMNPKATKII 420

```



Db	361	RMSVWEKCVASGMDNEFVAVTSTNAAKIFNYPKRGVAVGSDADLVINPRATKVF	420
Qy	421	SAKTHNLNVEYNIFEGVECRGAPAVVISQGRVALEDGKMFVTPCAGRFVPRKTFPDPVYK	480
Db	421	SAKSHNLNVEYNIFEGVECRGAPAVVISQGRVLEDCGVLVTPCAGRFVPRKTFPDPVYK	480
Qy	481	RIKARNLAEIHGVPRLGDPVHEVWVPAPKPGSGAPARASCPGKISVPPVRNLHSGGFS	540
Db	481	RIKARNLAEIHGVPRLGDPVHEVWVPAPKPGSGGTQARASCSGKISVPPVRNLHSGGFS	540
Qy	541	LSGSQADHDHARRTAQKIMAPPGGRSNTISLS	572
Db	541	LSGSQADHDHARRTAQKIMAPPGGRSNTISLS	572
RESULT 3			
DPY4 RAT	DPY4 RAT		
AC	Q62951;	STANDARD;	PRT; 564 AA.
DT	15-JUL-1998 (Rel. 36, Created)		
DT	15-JUL-1998 (Rel. 36, Last sequence update)		
DT	05-JUL-2004 (Rel. 44, Last annotation update)		
DE	Dihydropyrimidinase related protein-4 (DRP-4) (Collapsin response		
DE	mediator protein 3) (CRMP-3) (UNC33-like phosphoprotein 4) (ULIP4		
DE	protein) (Fragment).		
GN	Names=Dpye14; Synonyms=Crmp3, Ulip4;		
OS	Rattus norvegicus (Rat).		
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.		
OX	NCBI_TaxID=10116;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	TISSUE=Brain;		
RX	MSDLIN=96424532; PubMed=8815901;		
RA	Wang L.; Strittmatter S.M.;		
RA	"A family of rat CRMP genes is differentially expressed in the nervous		
RT	system."		
RL	J. Neurosci. 16:6197-6207(1996).		
CC	-1- FUNCTION: May be involved in the axonal growth during development		
CC	by transducing signals from different semaphorins.		
CC	-1- SUBCELLULAR LOCATION: Cytoplasmic (Potential).		
CC	-1- DEVELOPMENTAL STAGE: Expressed transiently in developing spinal		
CC	cord and selectively in the postnatal cerebellum.		
CC	-1- SIMILARITY: Belongs to the dehydropyrimidinase family.		
CC	-----		
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration		
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -		
CC	the European Bioinformatics Institute. There are no restrictions on its		
CC	use by non-profit institutions as long as its content is in no way		
CC	modified and this statement is not removed. Usage by and for commercial		
CC	entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a>		
CC	or send an email to <a href="mailto:license@isb-sib.ch">license@isb-sib.ch</a> ).		
CC	-----		
DR	EMBL; U52103; AB03281.1; -.		
DR	HSP; Q45515; IKID.		
DR	InterPro; IPR006680; Amidohydro_1.		
DR	InterPro; IPR005847; DHOase.		
DR	InterPro; IPR011059; Metallo_hydrolase.		
DR	Pfam; PF01979; Amidohydro_1; 1.		
DR	ProDom; PD000518; DHOase; 1.		
FT	NON TER		
SQ	SEQUENCE	564 AA; 61085 MW; 42050891CC1436D2 CRC64;	
Query Match			
Best Local Similarity 93.78; Score 2795; DB 1; Length 564;			
Matches 527; Conservative 21; Mismatches 16; Indels 0; Gaps 0;			
Qy	9	IPRITSDRLIRGRIWVNDQSFVADVHVEDGLIKQIGENLIVPGGKTTDAHGLVLP	68
Db	1	IPRITSDRLIRGRIWVNDQSFVADVHVEDGLIKQIGENLIVPGGKTTDAHGLVLP	60
Qy	69	GVDVHTRLQMPVLGMPADDFCQGTAAALAGGTTMLDHFVPPDTSVSLAAYEQWERRAD	128

```
DR ProDom; PD000518; DHOase; 1.
SQ SEQUENCE 574 AA; 62070 MM; 96E17AC0E96CASF4 CRC64;

Query Match      84.5%; Score 2519; DB 2; Length 574;
Best Local Similarity 81.4%; Pred. No. 9.1e-174;
Matches 467; Conservative 57; Mismatches 48; Indels 2; Gaps 1;

QY 1 MSFGQKKSIPRITSRLLIRGGRIVNDQSFYADVHVEDGLIKQIGENLIVPGGIKTIDA 60
DB 1 MSNPGKRSRGRTSEQLLIRGGKIIVNDQSFYADIVVEDGLIKQIGENLIVPGGRTIDA 60
QY 61 HGLMVLPGGVVHTRIQMPVLGTPADDFCQGTKAALAGGTTMILDHVPDPTGVSLLAAY 120
DB 61 YGQLVVPGGIDVHTRIQMAVMGASADDFVQGTTRAALAGGTTMIDHVLPEAGTSLAAY 120
QY 121 EQWRERADSAACDYSILHVDITWHSIEKELEALVKEKGVNSFLVFMAYKDRCCQSDSQ 180
DB 121 EQWRERADSRACCDYALHIDIPRWHSIEKELEALVKEKGVNSFLVFMAYKDRCCQSDSQ 180
QY 181 MYEIPSTIIRDLGALAOVHAENGDIIVEEQKRLLELIGITGPEGHVLSHPREVREAVYRAV 240
DB 181 MYEIFCIIIRDLGALAOVHAENGDIIVEEQKRLLELIGITGPEGHVLSHPREVREAVYRAI 240
QY 241 TIAQANCPLYVTKVMSKGAADAIAQAKRGVVVFGPEPTASLGTGSHYWSKNWAKAAA 300
DB 241 TIAQANCPLYVTKVMSKGAADAIAQAKRGVVVFGPEPTASLGTGSHYWSKNWAKAAA 300
QY 301 FVTSPPNPDPPTADHLTCLLSSGDLQVTSAGHCTTTAAQAVKGNFALIPGTTNGIEE 360
DB 301 FVTSPPNPDPPTADHLTCLLSSGDLQVTSAGHCTTTAAQAVKGNFALIPGTTNGIEE 360
QY 361 RSMWWEKCVASGKMDENFVATSTNAKIFNFPYRKGVRVAVGSDADLVINWPKATKII 420
DB 361 RMAIIWEKCVASGKMDENFVATSTNAKIFNFPYRKGVRVAVGSDADLVINWPKATKVI 420
QY 421 SAKTHNLNVEYNIIFEGVECRGAPAVVISQGRVLEDKGMFVTPGAGRFVPRKTFPDPVYK 480
DB 421 SAKTHNLNVEYNIIFEGVECRGAPAVVISQGRVLEDKGMFVTPGAGRFVPRKTFPDPVYK 480
QY 481 RIKARNLRLAIEHVGPRGLVDGPHVHVMVPAK--PGSGAPARASCPOKISVPPVPRNLHQSG 538
DB 481 RIKARNLRLAIEHVGPRGLVDGPHVHVMVPAK--PGSGAPARASCPOKISVPPVPRNLHQSG 540
QY 539 FSLSGSQADDDHARRTAQKIMAPPGSRNITSLS 572
DB 541 FSLSGSQADDDHARRTAQKIMAPPGSRNITSLS 574

RESULT 5
Q718G0 PRELIMINARY; PRT; 650 AA.
AC Q718G0
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DE Collapsin response mediator protein-3A.
GN Name=CRMP3A;
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22699123; PubMed=12814366;
RA Yuasa-Kawada J., Suzuki R., Kano F., Ohkawara T., Murata M., Noda M.;
RT "Axonal morphogenesis is controlled by antagonistic roles of two CRMP
RL subtypes in microtubule organization."
RL Eur. J. Neurosci. 17:2329-2343(2003).
DR EMBL; AF301551; A0414489.1; -
DR HSSP; P97427; 1KXK.
DR GO; GO:0016787; F:hydrolase activity; IEA.
DR InterPro; IPR006680; Amidohydro_1.
```

```
DR InterPro; IPR005847; DHOase.
DR InterPro; IPR011059; Metallo hydrolase.
DR Pfam; PF01979; Amidohydro_1.
DR ProDom; PD000518; DHOase; 1.
SQ SEQUENCE 650 AA; 69561 MM; E1FAB72370C26A4D CRC64;

Query Match      83.4%; Score 2488; DB 2; Length 650;
Best Local Similarity 81.8%; Pred. No. 1.9e-171;
Matches 459; Conservative 56; Mismatches 44; Indels 2; Gaps 1;

QY 14 SDRLLIRGRIVNDQSFYADVHVEDGLIKQIGENLIVPGGIKTIDAHGLMVLPGGVVH 73
DB 90 SEQLLIRGGKIIVNDQSFYADIVVEDGLIKQIGENLIVPGGRTVDAYGQLVVPGGIDVH 149
QY 74 TRLQMPVLGTPADDFCQGTKAALAGGTTMILDHVPDPTGVSLLAAYEOWRERADSAAC 133
DB 150 TRLQMAVMGASADDFVQGTTRAALAGGTTMIDHVLPEAGTSLAAYEOWRERADSRAC 209
QY 134 DYSLSHVDITRWHSIEKELEALVKEKGVNSFLVFMAYKDRCCQSDSQSYEIFSIRDLGA 193
DB 210 DYALHIDIPRWHSIEKELEALVKEKGVNSFLVFMAYKDRCCQSDSQSYEIFSIRDLGA 269
QY 194 LAQVHAENGDIIVEEQKRLLELIGITGPEGHVLSHPREVREAVYRAVTTAKQANCPLYVT 253
DB 270 LAQVHAENGDIIVEEQKRLLELIGITGPEGHVLSHPREVREAVYRAVTTAKQANCPLYVT 329
QY 254 KVMKGAADAIAQAKRGVVVFGPEPTASLGTGSHYWSKNWAKAAAFTVSPVNPDPPT 313
DB 330 KIMSKAADVVAQAKRGVVVFGPEPTASLGTGSHYWSKNWAKAAAFTVSPVNPDPPT 389
QY 314 ADHLTCLLSSGDLQVTSAGHCTTTAAQAVKGNFALIPGTTNGIEEWSMWEKCVASG 373
DB 390 PERLSLSSGDLQVAGSAHCTTTAAQAVKGNFALIPGTTNGIEEWSMWEKCVASG 449
QY 374 KMDNEFVATSTNAKIFNFPYRKGVRVAVGSDADLVINWPKATKIIISAKTHNLNVEYNI 433
DB 450 KMDNEFVATSTNAKIFNFPYRKGVRVAVGSDADLVINWPKATKIIISAKTHNLNVEYNI 509
QY 434 FSGVECRGAPAVVISQGRVLEDKGMFVTPGAGRFVPRKTFPDPVYKRIKARNLRLAIEH 493
DB 510 FSGTECHGAPAVVISQGRVLEDKGMFVTPGAGRFVPRKTFPDPVYKRIKARNLRLAIEH 569
QY 494 VPRGLYDGVHVMVPAK--PGSGAPARASCPOKISVPPVPRNLHQSGFSLSGSQADDDH 551
DB 570 VPRGLYDGVHVMVPAK--PGSGAPARASCPOKISVPPVPRNLHQSGFSLSGSQADDDH 629
QY 552 RRTAQKIMAPPGSRNITSLS 572
DB 630 RRTAQKIMAPPGSRNITSLS 650

RESULT 6
Q72TP7 PRELIMINARY; PRT; 668 AA.
AC Q72TP7
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE MGC53768 protein.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G., Schuler G.D.,
RA Klausen R.D., Collins F.S., Wagner K.H., Shenmen C.M., Schuler N.K.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore S.I., Wang J., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
```

RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E., Brownstein M.J., Uedlin T.B., Toshiyuki S., Carninci P., Prange C., Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J., Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H., Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W., Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A., Phney J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A., Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G., Blakesley R.W., Touchman J.W., Green E.D., Dickinson M.C., Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S., Kzywinski M.I., Skalska U., Smailus D.B., Schnerch A., Schein J.B., Jones S.J., Marra M.A.;  
RT "Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences";  
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Embryo;  
RX MEDLINE=22341132; PubMed=12454917; DOI=10.1002/dvdy.10174; Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W., EMBL; BC043880; AA043880.1; -;  
DR HSSP; P97427; IKCX.  
DR GO; GO:0016787; F:hydrolase activity; IEA.  
DR InterPro; IPR006680; Amidohydro\_1.  
DR InterPro; IPR005847; DHOase.  
DR InterPro; IPR011059; Metallo\_hydrolase.  
DR Pfam; PF01979; Amidohydro\_1; 1.  
DR ProDom; PD000518; DHOase; 1.  
SQ SEQUENCE 668 AA; 72485 MW; AB3994D680E3B500 CRC64;  
Query Match 81.9%; Score 2441.5; DB 2; Length 668;  
Best Local Similarity 79.5%; Pred. No. 4.7e-168;  
Matches 445; Conservative 67; Mismatches 47; Indels 1; Gaps 1;  
QY 14 SDRLIRGRVNDQSFYADVHVEDGLIKQIGENLVPGGIKTIDAGHLMVLPFGVDVH 73  
DB SDRLIRGGKI VNDQSFYADVHVEDGLIKQIGENLVPGGIKTIDAGHLMVLPFGVDVH 168  
QY 74 TRLOMPVLGTMTPADDFCQGTKAALAGGTTMILDHVPDPTGVSLLAAVEQWRERADSAAC 133  
DB TRLOMPVGMGTSADDFCQGTKAALAGGTTMILDHVPDPTGVSLLAAVEQWRERADSAAC 168  
QY 169 TRLOMPVGMGTSADDFCQGTKAALAGGTTMILDHVPDPTGVSLLAAVEQWRERADSAAC 228  
DB TRLOMPVGMGTSADDFCQGTKAALAGGTTMILDHVPDPTGVSLLAAVEQWRERADSAAC 228  
QY 134 DYSLSHVDITRWHSIKEELEALVKEGVNSFLVPMAYKDRCCQSDSOMVEIPSIIRDLGA 193  
DB DYSLSHVDITRWHSIKEELEALVKEGVNSFLVPMAYKDRCCQSDSOMVEIPSIIRDLGA 288  
QY 194 LAQVHAENGDI VBEQQRLEELGTGPEGHVLSHPEVRAEAVYRAVTIAKQANCPLYVT 253  
DB LAQVHAENGDI VBEQQRLEELGTGPEGHVLSHPEVRAEAVYRAVTIAKQANCPLYVT 348  
QY 254 KVMKGAADAIAQKRGRVVGEPITASIGTGSHYWSKNWAKAAAFVTSPPVNPDPPTT 313  
DB KVMKGAADAIAQKRGRVVGEPITASIGTGSHYWSKNWAKAAAFVTSPPVNPDPPTT 408  
QY 314 ADHLTCLSSGDLQVTSASCTTAAKAVGKONFALIPETNGIERMSWWEKCVASG 373  
DB ADHLTCLSSGDLQVTSASCTTAAKAVGKONFALIPETNGIERMSWWEKCVASG 468  
QY 409 PDHLTCLSSGDLQVTSASCTTAAKAVGKONFALIPETNGIERMSWWEKCVASG 468  
DB PDHLTCLSSGDLQVTSASCTTAAKAVGKONFALIPETNGIERMSWWEKCVASG 468  
QY 374 KMDENEFVATVSTNAAKIFNYPKRGVAVGSDADLVINPKATKIISAKTHLNVEYNI 433  
DB KMDENEFVATVSTNAAKIFNYPKRGVAVGSDADLVINPKATKIISAKTHLNVEYNI 528  
QY 469 KMDENEFVATVSTNAAKIFNYPKRGVAVGSDADLVINPKATKIISAKTHLNVEYNI 528  
DB KMDENEFVATVSTNAAKIFNYPKRGVAVGSDADLVINPKATKIISAKTHLNVEYNI 528  
QY 434 FEGVECRGAPAVVISQGRVALEPGKMFVTPGAGRFVPRKTPDPVYKRIKARNLAEIHG 493  
DB FEGVECRGAPAVVISQGRVALEPGKMFVTPGAGRFVPRKTPDPVYKRIKARNLAEIHG 493

Db 529 FEGMECYGAPEVVVISQGRVILEDNHVSQSGRFVPRKTPDPVYKRIKARNLAEIRG 588  
QY 494 VPRGLYDGPVHEVYVPAKPGSGAPA-RASCPGKISVPPVNLHQSFGSLSGSQAADHAR 552  
DB VPRGLYDGPVHEVYVPAKPGSGAPA-RASCPGKISVPPVNLHQSFGSLSGSQAADHAR 552  
QY 589 VPRGLYDGPVHEVYVPAKPGSGAPA-RASCPGKISVPPVNLHQSFGSLSGSQAADHAR 648  
DB VPRGLYDGPVHEVYVPAKPGSGAPA-RASCPGKISVPPVNLHQSFGSLSGSQAADHAR 648  
QY 553 RTAQKIMAPPGRSNTISLS 572  
DB RTAQKIMAPPGRSNTISLS 572  
QY 649 RTAQKIMAPPGRSNTISLS 668  
DB RTAQKIMAPPGRSNTISLS 668

RESULT 7  
DPY2\_CHICK STANDARD; PRT; 572 AA.  
ID QY0635;  
AC QY0635;  
DT 15-JUL-1998 (Rel. 36, Created)  
DT 15-JUL-1998 (Rel. 36, Last sequence update)  
DT 15-JUL-1998 (Rel. 36, Last annotation update)  
DE Dihydropyrimidinase related protein-2 (DRP-2) (Collapsin response mediator protein CRMP-62).  
OS Gallus gallus (Chicken).  
OC Rukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Archoauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;  
OC Gallus.  
OX NCBI\_TaxID=9031;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Dorsal root ganglion;  
RX MEDLINE=95364923; PubMed=7637782; DOI=10.1038/376509a0; Goshima Y., Nakamura F., Strittmatter P., Strittmatter S.M.;  
RA "Collapsin-induced growth cone collapse mediated by an intracellular protein related to UNC-33".  
RT Nature 376:509-514(1995).  
CC -!- FUNCTION: Involved in neuronal growth cone collapse.  
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (Potential).  
CC -!- SIMILARITY: Belongs to the dehydropyrimidinase family.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement. See <http://www.ebi.ac.uk/announcements> or send an email to [license@ebi.ac.uk](mailto:license@ebi.ac.uk).  
CC EMBL; U17277; AAA93200.1; -;  
DR FIR; S58889; S58889.  
DR HSSP; Q45515; IKID.  
DR InterPro; IPR006680; Amidohydro\_1.  
DR InterPro; IPR005847; DHOase.  
DR InterPro; IPR011059; Metallo\_hydrolase.  
DR Pfam; PF01979; Amidohydro\_1; 1.  
DR ProDom; PD000518; DHOase; 1.  
DR ProDom; PD000518; DHOase; 1.  
SQ SEQUENCE 572 AA; 62330 MW; 85DB9E3DD5E54D8D CRC64;  
Query Match 79.4%; Score 2368; DB 1; Length 572;  
Best Local Similarity 76.5%; Pred. No. 7.9e-163;  
Matches 437; Conservative 68; Mismatches 66; Indels 0; Gaps 0;  
QY 1 MSFGQKKSIPRITSRLIRGRVNDQSFYADVHVEDGLIKQIGENLVPGGIKTIDA 60  
DB MSFGQKKSIPRITSRLIRGRVNDQSFYADVHVEDGLIKQIGENLVPGGIKTIDA 60  
QY 61 HGLMVLPGVDVHTRLOMPVLGTMTPADDFCQGTKAALAGGTTMILDHVPDPTGVSLLAA 120  
DB HGLMVLPGVDVHTRLOMPVLGTMTPADDFCQGTKAALAGGTTMILDHVPDPTGVSLLAA 120  
QY 121 EQWRERADSAACDYSLSHVDITRWHSIKEELEALVKEGVNSFLVPMAYKDRCCQSDSQ 180  
DB EQWRERADSAACDYSLSHVDITRWHSIKEELEALVKEGVNSFLVPMAYKDRCCQSDSQ 180  
QY 121 DQWRERADSAACDYSLSHVDITRWHSIKEELEALVKEGVNSFLVPMAYKDRCCQSDSQ 180  
DB DQWRERADSAACDYSLSHVDITRWHSIKEELEALVKEGVNSFLVPMAYKDRCCQSDSQ 180  
QY 181 MYEIFSIRDLGALAQVHAENGDI VBEQQRLEELGTGPEGHVLSHPEVRAEAVYRAV 240

```

Db 181 IYEVLSVIRDIGATAQVHAENGDIIEBQQRILLELGTGPEGHVLSRPEVEAEAVNRAI 240
QY 241 TIAQANCPLYTVKWSKGAADIAIAQAKRGVVVFGPEPTASLGTGSHYWSKNWAKAAA 300
Db 241 TIANQNCPLYTVKWSKSAAEVIAQARKGTVVYGEPTASLGTGSHYWSKNWAKAAA 300
QY 301 FVTSPPVNDPTTADHLTCLSSGDLQVTSAGHCTFTTAQKAVGKDNFALIPEGTNGIEE 360
Db 301 FVTSPPVNDPTTDFLNSLLSCGDLQVTSAGHCTFTTAQKAVGKDNFALIPEGTNGIEE 360
QY 361 RMSMWKCVASGKMDENEFVAVTSTNAAKIFNFYPRKGRVAVGSDADLVINWPKATKII 420
Db 361 RMSIWDKAVVTKMDENQFVAVTSTNAAKIFNLYPRKGRIAVGSDADLVINWPKATKII 420
QY 421 SAKTHNLNVEYNIPEGVECRGAPAVVISQGRVALEDGKMFVTPGAGRFPVPRKTPFDPVYK 480
Db 421 SAKTHNISLEYNIFEGMECRGSLVVISQKIVLEDGNLHVTEGSGRYTPRKPFDPVYK 480
QY 481 RIKARNRLAEIHGVPRLGVDGVPVHVMPAKPGSGAPARASCPGKISVPPVRLNHOSGFS 540
Db 481 RIKARSLRLAELRGVPRGLVDGVPVCEVSVTPKTVTPASSAKTSPAKQQAPPVRLNHOSGFS 540
QY 541 LSGSQADDDHARRTAQKIMAPPGGRSNITSL 571
Db 541 LSGAQIDDDNIPTTQRIIVAPPGGRANITSL 571

RESULT 8
Q71SG1 PRELIMINARY; PRT; 572 AA.
AC Q71SG1
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Collapsin response mediator protein-2B.
GN Names=CRMP2B;
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22699123; PubMed=12814366;
RA Yuasa-Kawada J., Suzuki R., Kano F., Ohkawara T., Murata M., Noda M.;
RT "Axonal morphogenesis controlled by antagonistic roles of two CRMP
subtypes in microtubule organization.";
RL Eur. J. Neurosci. 17:2329-2343(2003).
DR EMBL; AF301550; AAQ14488.1; -.
DR HSSP; P97427; IKCX.
DR GO; GO:0016787; F:hydrolase activity; IEA.
DR InterPro; IPR006680; Amidohydro_1.
DR InterPro; IPR011550; Amidohydro_like.
DR InterPro; IPR005847; DHOase.
DR InterPro; IPR011059; Metallo_hydrolase.
DR Pfam; PF01979; Amidohydro_1; 1.
DR ProDom; PD001248; Amidohydro_like; 1.
DR ProDom; PD000518; DHOase; 1.
SQ SEQUENCE 572 AA; 62258 MW; 07DB9D888306F8DB CRC64;

Query Match 79.2%; Score 2361; DB 2; Length 572;
Best Local Similarity 76.4%; Pred. No. 2.5e-162;
Matches 436; Conservative 68; Mismatches 67; Indels 0; Gaps 0;

QY 1 MSFGKKSIPRITSDDLIRGGRIVNDQSFYADVHVEDGLIKIGENLIVPGGIKTIDA 60
Db 1 MSYQGGKNIPRITSDDLIRGGRIVNDQSFYADIYMEDGLIKIGENLIVPGGVKTIEA 60
QY 61 HGLMVLPGGVVDVHTRIQMPVLGMPVADDPFCQGTAAALAGTTTILDRHVPFDTGVSLAAY 120
Db 61 HGRMVIPGGIDVHTRFQMPQGGMTSADDPFQGTAAALAGTTTIIIDHVVPEPGTSLTAF 120

```

```

QY 121 EOWBRADSAACDYSLSHVDITRMHESIKEELEALVKEKGVNSFLVFMAYKDRCCQSDSQ 180
Db 121 DQWRWALWSKCCDYSLSHVDITRMHKGVEEALVKDHGVNSFLVYMAKDRPFQLSDSQ 180
QY 181 MYEIFSIIRDLGALAQVHAENGDIIEBQQRILLELGTGPEGHVLSHPEVEAEAVYRAV 240
Db 181 IYEVLSVIRDIGATAQVHAENGDIIEBQQRILLELGTGPEGHVLSRPEVEAEAVNRAI 240
QY 241 TIAQANCPLYTVKWSKGAADIAIAQAKRGVVVFGPEPTASLGTGSHYWSKNWAKAAA 300
Db 241 TIANQNCPLYTVKWSKSAAEVIAQARKGTVVYGEPTASLGTGSHYWSKNWAKAAA 300
QY 301 FVTSPPVNDPTTADHLTCLSSGDLQVTSAGHCTFTTAQKAVGKDNFALIPEGTNGIEE 360
Db 301 FVTSPPVNDPTTDFLNSLLSCGDLQVTSAGHCTFTTAQKAVGKDNFALIPEGTNGIEE 360
QY 361 RMSMWKCVASGKMDENEFVAVTSTNAAKIFNFYPRKGRVAVGSDADLVINWPKATKII 420
Db 361 RMSIWDKAVVTKMDENQFVAVTSTNAAKIFNLYPRKGRIAVGSDADLVINWPKATKII 420
QY 421 SAKTHNLNVEYNIPEGVECRGAPAVVISQGRVALEDGKMFVTPGAGRFPVPRKTPFDPVYK 480
Db 421 SAKTHNISLEYNIFEGMECRGSLVVISQKIVLEDGNLHVTEGSGRYTPRKPFDPVYK 480
QY 481 RIKARNRLAEIHGVPRLGVDGVPVHVMPAKPGSGAPARASCPGKISVPPVRLNHOSGFS 540
Db 481 RIKARSLRLAELRGVPRGLVDGVPVCEVSVTPKTVTPASSAKTSPAKQQAPPVRLNHOSGFS 540
QY 541 LSGSQADDDHARRTAQKIMAPPGGRSNITSL 571
Db 541 LSGAQIDDDNIPTTQRIIVAPPGGRANITSL 571

RESULT 9
DPY2_BOVIN STANDARD; PRT; 572 AA.
AC O02675;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Dihydropyrimidinase related protein-2 (DRP-2) (Neural specific protein
DE NSP60).
GN Name=DPYSL2;
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Ruminantia; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RA Kamata T.K.;
RL Submitted (APR-1997) to the EMBL/GenBank/DBJ databases.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (Potential).
CC -!- SIMILARITY: Belongs to the dehydropyrimidinase family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; U83278; AAB80618.1; -.
DR HSSP; Q45515; 1KID.
DR InterPro; IPR006680; Amidohydro_1.
DR InterPro; IPR005847; DHOase.
DR InterPro; IPR011059; Metallo_hydrolase.
DR Pfam; PF01979; Amidohydro_1; 1.
DR ProDom; PD001248; Amidohydro_like; 1.
DR ProDom; PD000518; DHOase; 1.
SQ SEQUENCE 572 AA; 62277 MW; 343507ACB9D91BDE CRC64;

```

Query Match		78.9%; Score 2353; DB 1; Length 572;
Best Local Similarity		75.8%; Pred. No. 9.6e-162;
Matches 433; Conservative 72; Mismatches 66; Indels 0; Gaps 0;		
Qy	1	MSFGKKSIPRTSDRLIRGRVNDQSFYADVHVEDGLIKQIGENLVPGGIKTIDA 60
Db	1	MSVQGGKNIPIRTSDRLIRGGKIVNDQSFYADIYMEDGLIKQIGENLVPGGVKTIEA 60
Qy	61	HGLMVLPGGVVHTRLOMPVLGMPADDFQCGTKAALAGGTTMLDHPVDTGVSLLAA 120
Db	61	HSRMVPGGIDVHTRFQMPDQGMTSADDFQGTKAALAGGTTMLDHPVDPGTSLLAA 120
Qy	121	EQWRERADSAACDYSLSHVDITRWHSIKBELEALVKEGVNSFLVPMAYKDRCCQSDSQ 180
Db	121	DQWRERADSAACDYSLSHVDITRWHSIKBELEALVKEGVNSFLVPMAYKDRFQLTDSQ 180
Qy	181	MYEIPSIIRDLGALAOVHAENGDIYVEEQKRLLELGLITGPEGHVLSHPREVEAEVYRAV 240
Db	181	IYEVLSVIRDIGALAOVHAENGDIYAEEOQRILDLGITGPEGHVLSRPEVEAEAVNRSI 240
Qy	241	TIKQANCPLYVTKWSKGAADAIQAQRKRGVVVFGPEPITASLGTDGSHYWSKNWAKAAA 300
Db	241	TIANTQNCPLIYTKWSKSAAEVIAQAQRKGTVVYGEPIITASLGTDGSHYWSKNWAKAAA 300
Qy	301	FVTSPPVNPDPPTADHLTCLLSGDLQVTSAGHCTFTTAQKAVGKONFALIPGTNGIEE 360
Db	301	FVTSPLSPDPPTDFLNSLLSCGDLQVTSAGHCTFTTAQKAVGKONFTLIPGTNGTEE 360
Qy	361	RMSVMWEKCVASGKMDNEFVAVTSTNAAKIFNYPKRGVAVGSDADLVINPNKATKII 420
Db	361	RMSVWDKAVVTGMDNEFVAVTSTNAAKFVNLVPRKGRISVSGSDADLVINPDSVKTI 420
Qy	481	RIKARNRLABIHGVRGLYDGPVHVWPAKPGSGAPARASCPCGKISVPPVRLHSGGFS 540
Db	481	RIKARSLAELRGVPRGLYDGPVCEVSVTPKTVTTPASSAKTSPAKQOAPPVRLHSGGFS 540
Qy	541	LSGSQADDDHARTAQKIMAPPGGRANITSL 571
Db	541	LSGAQIDDDNIPRRTTORIVAPPGGRANITSL 571
RESULT 10		
ID	DPY2_RAT	
AC	P47942;	STANDARD; PRT; 572 AA.
DT	01-FEB-1996 (Rel. 33, Created)	
DT	01-FEB-1996 (Rel. 33, Last sequence update)	
DT	05-JUL-2004 (Rel. 44, Last annotation update)	
DE	Dihydropyrimidinase related protein-2 (DRP-2) (Turned on after	
DE	division; 64 kDa protein) (TOAD-64) (Collapsin response mediator	
DE	protein 2) (CRMP-2).	
GN	Names=Dpy812;	
OS	Rattus norvegicus (Rat).	
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;	
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.	
OX	NCBI_TaxID=101116;	
RN	SEQUENCE FROM N.A., AND SEQUENCE OF 136-142; 402-418; 441-450 AND	
RP	499-511;	
RC	STRAIN=Sprague-Dawley; TISSUE=Brain;	
RA	MEDLINE=J6033765; PubMed=7472434;	
RA	Minturn J.E., Fryer H.J.L., Geschwind D.H., Hockfield S.;	
RT	"TOAD-64", a gene expressed early in neuronal differentiation in the	
RT	rat, is related to unc-33, a C. elegans gene involved in axon	
RT	outgrowth.";	
RL	J. Neurosci. 15:6757-6766(1995).	
CC	-!- FUNCTION: May have a role in axon elaboration.	
CC	-!- SUBCELLULAR LOCATION: Tightlly, but noncovalently, associated with	
CC	membranes.	
Query Match		78.8%; Score 2349; DB 1; Length 572;
Best Local Similarity		76.0%; Pred. No. 1.9e-161;
Matches 434; Conservative 70; Mismatches 67; Indels 0; Gaps 0;		
Qy	1	MSFGKKSIPRTSDRLIRGRVNDQSFYADVHVEDGLIKQIGENLVPGGIKTIDA 60
Db	1	MSVQGGKNIPIRTSDRLIRGGKIVNDQSFYADIYMEDGLIKQIGENLVPGGVKTIEA 60
Qy	61	HGLMVLPGGVVHTRLOMPVLGMPADDFQCGTKAALAGGTTMLDHPVDTGVSLLAA 120
Db	61	HSRMVPGGIDVHTRFQMPDQGMTSADDFQGTKAALAGGTTMLDHPVDPGTSLLAA 120
Qy	121	EQWRERADSAACDYSLSHVDITRWHSIKBELEALVKEGVNSFLVPMAYKDRCCQSDSQ 180
Db	121	DQWRERADSAACDYSLSHVDITRWHSIKBELEALVKEGVNSFLVPMAYKDRFQLTDSQ 180
Qy	181	MYEIPSIIRDLGALAOVHAENGDIYVEEQKRLLELGLITGPEGHVLSHPREVEAEVYRAV 240
Db	181	IYEVLSVIRDIGALAOVHAENGDIYAEEOQRILDLGITGPEGHVLSRPEVEAEAVNRSI 240
Qy	241	TIKQANCPLYVTKWSKGAADAIQAQRKRGVVVFGPEPITASLGTDGSHYWSKNWAKAAA 300
Db	241	TIANTQNCPLIYTKWSKSAAEVIAQAQRKGTVVYGEPIITASLGTDGSHYWSKNWAKAAA 300
Qy	301	FVTSPPVNPDPPTADHLTCLLSGDLQVTSAGHCTFTTAQKAVGKONFALIPGTNGIEE 360
Db	301	FVTSPLSPDPPTDFLNSLLSCGDLQVTSAGHCTFTTAQKAVGKONFTLIPGTNGTEE 360
Qy	361	RMSVMWEKCVASGKMDNEFVAVTSTNAAKIFNYPKRGVAVGSDADLVINPNKATKII 420
Db	361	RMSVWDKAVVTGMDNEFVAVTSTNAAKFVNLVPRKGRISVSGSDADLVINPDSVKTI 420
Qy	421	SAKTHNLNVEYNIPEGVECRGAPAVVISQGRVALEDGKMFVTPGAGRFPVPRKTFPPFVYK 480
Db	421	SAKTHNSLEYNIPEGMECRGSLPVVISQGVLEDTGLHVTGSGRYIPRKTFPPFVYK 480
Qy	481	RIKARNRLABIHGVRGLYDGPVHVWPAKPGSGAPARASCPCGKISVPPVRLHSGGFS 540
Db	481	RIKARSLAELRGVPRGLYDGPVCEVSVTPKTVTTPASSAKTSPAKQOAPPVRLHSGGFS 540
Qy	541	LSGSQADDDHARTAQKIMAPPGGRANITSL 571
Db	541	LSGAQIDDDNIPRRTTORIVAPPGGRANITSL 571
RESULT 11		
ID	Q6P5D0	PRELIMINARY; PRT; 572 AA.
Query Match		78.8%; Score 2349; DB 1; Length 572;
Best Local Similarity		76.0%; Pred. No. 1.9e-161;
Matches 434; Conservative 70; Mismatches 67; Indels 0; Gaps 0;		
Qy	1	MSFGKKSIPRTSDRLIRGRVNDQSFYADVHVEDGLIKQIGENLVPGGIKTIDA 60
Db	1	MSVQGGKNIPIRTSDRLIRGGKIVNDQSFYADIYMEDGLIKQIGENLVPGGVKTIEA 60
Qy	61	HGLMVLPGGVVHTRLOMPVLGMPADDFQCGTKAALAGGTTMLDHPVDTGVSLLAA 120
Db	61	HSRMVPGGIDVHTRFQMPDQGMTSADDFQGTKAALAGGTTMLDHPVDPGTSLLAA 120
Qy	121	EQWRERADSAACDYSLSHVDITRWHSIKBELEALVKEGVNSFLVPMAYKDRCCQSDSQ 180
Db	121	DQWRERADSAACDYSLSHVDITRWHSIKBELEALVKEGVNSFLVPMAYKDRFQLTDSQ 180
Qy	181	MYEIPSIIRDLGALAOVHAENGDIYVEEQKRLLELGLITGPEGHVLSHPREVEAEVYRAV 240
Db	181	IYEVLSVIRDIGALAOVHAENGDIYAEEOQRILDLGITGPEGHVLSRPEVEAEAVNRSI 240
Qy	241	TIKQANCPLYVTKWSKGAADAIQAQRKRGVVVFGPEPITASLGTDGSHYWSKNWAKAAA 300
Db	241	TIANTQNCPLIYTKWSKSAAEVIAQAQRKGTVVYGEPIITASLGTDGSHYWSKNWAKAAA 300
Qy	301	FVTSPPVNPDPPTADHLTCLLSGDLQVTSAGHCTFTTAQKAVGKONFALIPGTNGIEE 360
Db	301	FVTSPLSPDPPTDFLNSLLSCGDLQVTSAGHCTFTTAQKAVGKONFTLIPGTNGTEE 360
Qy	361	RMSVMWEKCVASGKMDNEFVAVTSTNAAKIFNYPKRGVAVGSDADLVINPNKATKII 420
Db	361	RMSVWDKAVVTGMDNEFVAVTSTNAAKFVNLVPRKGRISVSGSDADLVINPDSVKTI 420
Qy	421	SAKTHNLNVEYNIPEGVECRGAPAVVISQGRVALEDGKMFVTPGAGRFPVPRKTFPPFVYK 480
Db	421	SAKTHNSLEYNIPEGMECRGSLPVVISQGVLEDTGLHVTGSGRYIPRKTFPPFVYK 480
Qy	481	RIKARNRLABIHGVRGLYDGPVHVWPAKPGSGAPARASCPCGKISVPPVRLHSGGFS 540
Db	481	RIKARSLAELRGVPRGLYDGPVCEVSVTPKTVTTPASSAKTSPAKQOAPPVRLHSGGFS 540
Qy	541	LSGSQADDDHARTAQKIMAPPGGRANITSL 571
Db	541	LSGAQIDDDNIPRRTTORIVAPPGGRANITSL 571
RESULT 11		
ID	Q6P5D0	PRELIMINARY; PRT; 572 AA.

CC -!- DEVELOPMENTAL STAGE: Expressed immediately after neuronal birth  
 CC and is dramatically down-regulated in the adult.  
 CC -!- SIMILARITY: Belongs to the dehydroprimidinae family.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.ebi.ac.uk/announcements/>  
 CC or send an email to [license@ebi.ac.uk](mailto:license@ebi.ac.uk)).

CC EMBL; Z46882; CAA86981.1; -  
 DR PIR; A59280; S49985.  
 DR HSSP; Q45515; 1K1D.  
 DR RGD; 2517; Dpy812.  
 DR InterPro; IPR006680; Amidohydro\_1.  
 DR InterPro; IPR005847; DHOase.  
 DR InterPro; IPR011059; Metallo\_hydrolase.  
 DR Pfam; PF01979; Amidohydro\_1; 1.  
 DR ProDom; PD001248; Amidohydro\_like; 1.  
 DR ProDom; PD000518; DHOase; 1.  
 KW Direct protein sequencing.  
 SQ SEQUENCE 572 AA; 62277 MW; C031P3BC039AA737 CRC64;

Query Match 78.8%; Score 2349; DB 1; Length 572;  
 Best Local Similarity 76.0%; Pred. No. 1.9e-161;  
 Matches 434; Conservative 70; Mismatches 67; Indels 0; Gaps 0;

Qy 1 MSFGKKSIPRTSDRLIRGRVNDQSFYADVHVEDGLIKQIGENLVPGGIKTIDA 60  
 Db 1 MSVQGGKNIPIRTSDRLIRGGKIVNDQSFYADIYMEDGLIKQIGENLVPGGVKTIEA 60  
 Qy 61 HGLMVLPGGVVHTRLOMPVLGMPADDFQCGTKAALAGGTTMLDHPVDTGVSLLAA 120  
 Db 61 HSRMVPGGIDVHTRFQMPDQGMTSADDFQGTKAALAGGTTMLDHPVDPGTSLLAA 120  
 Qy 121 EQWRERADSAACDYSLSHVDITRWHSIKBELEALVKEGVNSFLVPMAYKDRCCQSDSQ 180  
 Db 121 DQWRERADSAACDYSLSHVDITRWHSIKBELEALVKEGVNSFLVPMAYKDRFQLTDSQ 180  
 Qy 181 MYEIPSIIRDLGALAOVHAENGDIYVEEQKRLLELGLITGPEGHVLSHPREVEAEVYRAV 240  
 Db 181 IYEVLSVIRDIGALAOVHAENGDIYAEEOQRILDLGITGPEGHVLSRPEVEAEAVNRSI 240  
 Qy 241 TIKQANCPLYVTKWSKGAADAIQAQRKRGVVVFGPEPITASLGTDGSHYWSKNWAKAAA 300  
 Db 241 TIANTQNCPLIYTKWSKSAAEVIAQAQRKGTVVYGEPIITASLGTDGSHYWSKNWAKAAA 300  
 Qy 301 FVTSPPVNPDPPTADHLTCLLSGDLQVTSAGHCTFTTAQKAVGKONFALIPGTNGIEE 360  
 Db 301 FVTSPLSPDPPTDFLNSLLSCGDLQVTSAGHCTFTTAQKAVGKONFTLIPGTNGTEE 360  
 Qy 361 RMSVMWEKCVASGKMDNEFVAVTSTNAAKIFNYPKRGVAVGSDADLVINPNKATKII 420  
 Db 361 RMSVWDKAVVTGMDNEFVAVTSTNAAKFVNLVPRKGRISVSGSDADLVINPDSVKTI 420  
 Qy 421 SAKTHNLNVEYNIPEGVECRGAPAVVISQGRVALEDGKMFVTPGAGRFPVPRKTFPPFVYK 480  
 Db 421 SAKTHNSLEYNIPEGMECRGSLPVVISQGVLEDTGLHVTGSGRYIPRKTFPPFVYK 480  
 Qy 481 RIKARNRLABIHGVRGLYDGPVHVWPAKPGSGAPARASCPCGKISVPPVRLHSGGFS 540  
 Db 481 RIKARSLAELRGVPRGLYDGPVCEVSVTPKTVTTPASSAKTSPAKQOAPPVRLHSGGFS 540  
 Qy 541 LSGSQADDDHARTAQKIMAPPGGRANITSL 571  
 Db 541 LSGAQIDDDNIPRRTTORIVAPPGGRANITSL 571

RESULT 11  
 Q6P5D0  
 ID Q6P5D0 PRELIMINARY; PRT; 572 AA.







```
DR Pfam; PF01979; Amidohydro 1; 1.
DR ProDom; PD001248; Amidohydro 1; 1.
DR ProDom; PD000518; DHOase; 1.
SQ SEQUENCE 572 AA; 62170 MW; 7FDAE8A242ACF62D CRC64;

Query Match
Best Local Similarity 78.2%; Score 2331; DB 1; Length 572;
Matches 431; Conservative 70; Mismatches 70; Indels 0; Gaps 0;

QY 1 MSFQKKSIPIRISDRLLIRGRVNDQSFYADVHVDEGLIKQIGENLIVPGGIKTIDA 60
Db 1 MSYQKKNIPIRISDRLLIRGRVNDQSFYADVHVDEGLIKQIGENLIVPGGIKTIDA 60
QY 61 HGLMVLPGGVVHTRLQMPVLGMPADDFCOGTTAAALAGGTTMLDHRVFPDPTGVSLAA 120
Db 61 HSRMVLPGGVVHTRLQMPVLGMPADDFCOGTTAAALAGGTTMLDHRVFPDPTGVSLAA 120
QY 121 EQWREADSAACDYSLSHVDITRWHSIKLEALVKEGVNSFLVFMAYKDRCCQSDSQ 180
Db 121 DQWREADSAACDYSLSHVDITRWHSIKLEALVKEGVNSFLVFMAYKDRCCQSDSQ 180
QY 181 MYEIFSIRDLGALAQAHAENGDI VEEBQKRLLELGTGPEGHVLSPHEVEAEAVYRA 240
Db 181 IYEVLSVIRDIGAIAQVHAENGDI VEEBQKRLLELGTGPEGHVLSPHEVEAEAVYRA 240
QY 241 TIAKQANCLVYTKVMSKGAADAIQAQRGVVVFGEPI TASIGTDSHYNSKWKAKAA 300
Db 241 TIANQNCPLVYTKVMSKGAADAIQAQRGVVVFGEPI TASIGTDSHYNSKWKAKAA 300
QY 301 FVTSPPVNPDPPTADHLTCLLSGGDLQVTSAGHCTFTTAQKAVGKDNFALIPGTINGIE 360
Db 301 FVTSPPVNPDPPTADHLTCLLSGGDLQVTSAGHCTFTTAQKAVGKDNFALIPGTINGIE 360
QY 361 RMSWWEKCVASGKNDENEFVATSTNAKIFNFPYPRKGRVAVGSDADLVWNPATKII 420
Db 361 RMSVWIDKAVVTGKNDENQFVATSTNAKIFNFPYPRKGRVAVGSDADLVWNPATKII 420
QY 421 SAKTHNLNVEYNIPEGVCRGAPAVVISQGRVALEDGKMFVTPGAGRFVPRKTPDPVYK 480
Db 421 SAKTHNSALEYNIPEGVCRGAPAVVISQGRVALEDGKMFVTPGAGRFVPRKTPDPVYK 480
QY 481 RIKARNLRLAEIHGVPRLGIDGVPVHEVMVPAKPGSGAPARASCPGKISVPPVRLNHQSGFS 540
Db 481 RIKARSLRLAEIHGVPRLGIDGVPVHEVMVPAKPGSGAPARASCPGKISVPPVRLNHQSGFS 540
QY 541 LSGSQADDDHIARTAKIMAPPGRSNTSL 571
Db 541 LSGAQIDDDNIPRRTTQRIVAPPGGRANITSL 571

RESULT 14
Q71SG2 PRELIMINARY; PRT; 676 AA.
ID Q71SG2
AC Q71SG2;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Collapsin response mediator protein-2A.
GN Name=CRMP2A;
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22699123; PubMed=12814366;
RA Yuasa-Kawada J., Suzuki R., Kano F., Ohkawara T., Murata M., Noda M.;
RT "Axonal morphogenesis controlled by antagonistic roles of two CRMP
RL subtypes in microtubule organization.";
DR Eur. J. Neurosci. 17:2329-2343 (2003).
DR EMBL; AF301549; AAK14487.1; -.
DR HSSP; P97427; 1KCY.

DR GO; GO:0016787; F:hydrolase activity; IEA.
DR InterPro; IPR006680; Amidohydro 1.
DR InterPro; IPR011550; Amidohydro 1.
DR InterPro; IPR005847; DHOase.
DR Pfam; PF01979; Amidohydro 1; 1.
DR ProDom; PD001248; Amidohydro 1; 1.
DR ProDom; PD000518; DHOase; 1.
SQ SEQUENCE 676 AA; 73440 MW; 09F6404D0943856A CRC64;

Query Match
Best Local Similarity 77.6%; Score 2315; DB 2; Length 676;
Matches 429; Conservative 70; Mismatches 71; Indels 2; Gaps 1;

QY 2 SFQKKSIPIRIT--SDRLIRGRVNDQSFYADVHVDEGLIKQIGENLIVPGGIKTID 59
Db 104 SATGKEALQNLDKSDRLIKGKI VNDQSFYADYIMEDGLIKQIGENLIVPGGIKTIE 163
QY 60 AHGLMVLPGGVVHTRLQMPVLGMPADDFCOGTTAAALAGGTTMLDHRVFPDPTGVSLAA 119
Db 164 AGRMVLPGGVVHTRLQMPVLGMPADDFCOGTTAAALAGGTTMLDHRVFPDPTGVSLAA 223
QY 120 YEQWREADSAACDYSLSHVDITRWHSIKLEALVKEGVNSFLVFMAYKDRCCQSDS 179
Db 224 FQWREADSAACDYSLSHVDITRWHSIKLEALVKEGVNSFLVFMAYKDRCCQSDS 283
QY 180 QMYEIFSIRDLGALAQAHAENGDI VEEBQKRLLELGTGPEGHVLSPHEVEAEAVYRA 239
Db 284 QIYEVLSVIRDIGAIAQVHAENGDI VEEBQKRLLELGTGPEGHVLSPHEVEAEAVYRA 343
QY 240 VTIKQANCLVYTKVMSKGAADAIQAQRGVVVFGEPI TASIGTDSHYNSKWKAKAA 299
Db 344 ITIANTNCPLVYTKVMSKGAADAIQAQRGVVVFGEPI TASIGTDSHYNSKWKAKAA 403
QY 300 AFVTSPPVNPDPPTADHLTCLLSGGDLQVTSAGHCTFTTAQKAVGKDNFALIPGTINGIE 359
Db 404 AFVTSPPVNPDPPTADHLTCLLSGGDLQVTSAGHCTFTTAQKAVGKDNFALIPGTINGIE 463
QY 360 RMSWWEKCVASGKNDENEFVATSTNAKIFNFPYPRKGRVAVGSDADLVWNPATKII 419
Db 464 RMSIWIWDKAVVTGKNDENQFVATSTNAKIFNFPYPRKGRVAVGSDADLVWNPATKII 523
QY 420 ISAKTHNLNVEYNIPEGVCRGAPAVVISQGRVALEDGKMFVTPGAGRFVPRKTPDPVYK 479
Db 524 ISAKTHNLSLEYNIEPEGVCRGAPAVVISQGRVALEDGKMFVTPGAGRFVPRKTPDPVYK 583
QY 480 KRIKARNLRLAEIHGVPRLGIDGVPVHEVMVPAKPGSGAPARASCPGKISVPPVRLNHQSGF 539
Db 584 KRIKARSLRLAEIHGVPRLGIDGVPVHEVMVPAKPGSGAPARASCPGKISVPPVRLNHQSGF 643
QY 540 LSGSQADDDHIARTAKIMAPPGRSNTSL 571
Db 644 LSGAQIDDDNIPRRTTQRIVAPPGGRANITSL 575

RESULT 15
Q6PB02 PRELIMINARY; PRT; 573 AA.
ID Q6PB02
AC Q6PB02;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE MGC68668 protein.
GN Name=MGC68668;
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RX TISSUE=Kidney;
DR MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
```



**THIS PAGE BLANK (USPTO)**

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 24, 2005, 15:57:51 ; Search time 272 Seconds  
(without alignments)  
202.338 Million cell updates/sec

Title: US-09-367-496C-8  
Perfect score: 2982  
Sequence: 1 MSFQGGKSPRITSDRLLR.....RTAQKIMAPPGRSNTSLIS 572

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR 79:\*  
1: pir1:\*  
2: pir2:\*  
3: pir3:\*  
4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB ID	Description
1	2368	79.4	572	2 S58889	collapsin response
2	2349	78.8	572	2 S49985	dihydropyrimidinas
3	2345	78.6	572	2 JC5317	dihydropyrimidinas
4	2193	73.5	570	2 S55525	Ulip protein - mou
5	2192	73.5	570	2 JC5318	dihydropyrimidinas
6	2157	72.3	572	2 JC5316	dihydropyrimidinas
7	1556	52.2	519	2 JC5315	dihydropyrimidinas
8	1550	52.0	519	2 S70581	dihydropyrimidinas
9	1246	41.8	520	2 T20007	hypothetical prote
10	1123	37.7	544	2 T23968	hypothetical prote
11	1083	36.3	489	2 AF3457	dihydropyrimidinas
12	1074.5	36.0	479	2 H83590	dihydropyrimidinas
13	1031.5	34.6	485	2 AH2869	dihydropyrimidinas
14	1031.5	34.6	506	2 B97646	dihydropyrimidinas
15	838	28.1	471	1 JC2310	dihydropyrimidinas
16	782.5	26.2	854	2 S33558	unc-33 protein - C
17	692.5	23.2	467	2 T28685	hypothetical prote
18	673	22.6	465	2 A65071	hypothetical prote
19	673	22.6	465	2 B91097	hypothetical prote
20	673	22.6	465	2 F85942	hypothetical prote
21	335.5	11.3	433	2 AE0567	probable allantoin
22	327	11.0	444	2 C87297	dihydroorotase [im
23	326.5	10.9	448	2 E75429	probable allantoin
24	318	10.7	438	2 E83938	allantoinase BH230
25	316.5	10.6	453	1 G64782	probable allantoin
26	296.5	9.9	430	2 B70959	probable dihydroor
27	296.5	9.9	446	1 C70016	allantoinase homol
28	296.5	9.9	454	2 A69017	dihydroorotase - M
29	284.5	9.5	496	2 AB2432	D-hydantoinase [im

30	277	9.3	478	2 AC3412	dihydroorotase (EC
31	260.5	8.7	422	2 D84403	dihydroorotase [im
32	249.5	8.4	424	2 D89892	dihydroorotase [im
33	249	8.4	424	2 H96963	dihydroorotase [im
34	243	8.1	430	2 T45151	dihydroorotase (EC
35	242.5	8.1	422	2 C70370	dihydroorotase - A
36	242.5	8.1	422	2 C95135	dihydroorotase, mu
37	242.5	8.1	426	2 AB1676	dihydroorotase hom
38	241.5	8.1	430	2 E86975	probable dihydroor
39	238.5	8.0	426	2 AE1304	dihydroorotase hom
40	235.5	7.9	422	2 E98003	dihydroorotase (EC
41	234	7.8	296	2 E90700	allantoinase [impo
42	234	7.8	296	2 H85550	allantoinase [impo
43	231.5	7.8	2198	2 T20371	hypothetical prote
44	229.5	7.7	2242	2 A57541	pyrimidine synthe
45	224.5	7.5	428	2 B83967	dihydroorotase pyr

ALIGNMENTS

RESULT 1

S58889  
collapsin response mediator protein, 62K - chicken  
C/Species: Gallus gallus (chicken)  
C/Date: 15-Feb-1996 #sequence\_revision 01-Mar-1996 #text\_change 09-Jul-2004  
C/Accession: S58889  
R/Goshima, Y.; Nakamura, F.; Strittmatter, P.; Strittmatter, S.M.  
Nature 376, 509-514, 1995  
A/Title: Collapsin-induced growth cone collapse mediated by an intracellular protein rel  
A/Reference number: S58889, MUID:95364923; PMID:7637782  
A/Accession: S58889  
A/Status: preliminary  
A/Molecule type: mRNA  
A/Residues: 1-572 <GS>  
A/Cross-references: UNIPROT:Q90635; EMBL:U17277; NID:g882146; PIDN:AAA93200.1; PID:g88214  
C/Superfamily: allantoinase; Bacillus dihydroorotase homology

Query Match 79.4%; Score 2368; DB 2; Length 572;  
Best Local Similarity 76.5%; Pred. No. 2.2e-164;  
Matches 437; Conservative 68; Mismatches 66; Indels 0; Gaps 0;

QY	1	MSFQGGKSPRITSDRLLRGRVNDQSFVADVHVEDGLIKQIGENLVPGGIKTIDA	60
DB	1	MSYQGGKNIPRITSDRLLRIGKGIKGVNDQSFVADVHVEDGLIKQIGENLVPGGVKTIEA	60
QY	61	HGLMVLPGGVVHTRLQMPVLGWTADDFCQGTAKALAGTTMILDHVPDGTGVSLAAY	120
DB	61	HGRMVIPGGIDVHTRFQMPQGGTADDFCQGTAKALAGTTMIIIDHVVEPGTGLLTAF	120
QY	121	EQWRERADSAACDYSLSHYDITRWHSIKBELEALVKEKGVNSFLVFMAYKDCQCSDSQ	180
DB	121	DQWRWADSKSCDYSLSHYDITRWHSIKBELEALVKEKGVNSFLVFMAYKDCQCSDSQ	180
QY	181	MYEIPSIIRDLAGALQVHAENGDIYVEEQKRLLELGITGPEGHVLSHPREVEAAVYRAV	240
DB	181	IYEVLVIRDIGATAQVHAENGDIYAEQQRILELGITGPEGHVLSHPREVEAAVNRAI	240
QY	241	TTAKQANCPLYTKVTKMSKGAADAIAQAKRGVVVFGEPTASLGTDGSHYKSNKAKAAA	300
DB	241	TTANQNTCLYITKVMSKSAAEVIAQAKRGVTVYGEPTASLGTDGSHYKSNKAKAAA	300
QY	301	FVTSPPVNPDPPTADHLTCLSSGDLQVTSAGHCTFTTAQKAVGKNFALIPGTTNGIEE	360
DB	301	FVTSPPVNPDPPTADHLTCLSSGDLQVTSAGHCTFTTAQKAVGKNFALIPGTTNGIEE	360
QY	361	RMSMWKCVASGKMDNEFVAVTSTNAAKIFNFPYPRKGRVAVGSDADLVINWPKATKII	420
DB	361	RMSIWDKAVVTGMDNQFVAVTSTNAAKIFNLYPRKGRVAVGSDADLVINWPKATKII	420
QY	421	SAKTHNLNVEYNIFEGVECRGAPAVVISQGRVALEKGMFTVTCAGRFVPRKTFPDPVYK	480
DB	421	SAKTHNISLEYNIFEGMECRGSLPVVISQGIKVLNHLVTEGSGRIYIPKPFPPDPVYK	480

Qy	481	RIKARNRLAEIHGVPRLGYDGPVHEVMVPKPGSGAPARASCFOKISVPPVRNLHQSGFS	540
Db	481	RIKARSRLAELRGVPRGLYDGPVCEVSVTKVTTPASSAKTSPAKQQAQPPVRNLHQSGFS	540
Qy	541	LSGSQADDDHIAARRTAQKIMAPPGRSNTSL	571
Db	541	LSGAQIDDNIPRETTQRIVAPPGRANITSL	571

RESULT 2

849985

diHydroxyrimidine-related protein 2 [similarity] - rat  
N/Alternate names: collapsin response mediator protein 2; TOAD-64  
C/Species: Rattus norvegicus (Norway rat)  
C/Date: 13-Jan-1995 #sequence\_revision 12-May-1995 #text\_change 09-Jul-2004  
C/Accession: A59280, S49985  
R/Minturn, J.B.; Fryer, H.J.; Geschwind, D.H.; Hockfield, S.  
J. Neurosci. 15, 6757-6766, 1995  
A/Title: TOAD-64, a gene expressed early in neuronal differentiation in the rat, is related to the human TOAD-64 gene  
A/Reference number: A59280; MUID:56033765; PMID:7472434  
A/Accession: A59280  
A/Status: preliminary; not compared with conceptual translation  
A/Molecule type: mRNA  
A/Residues: 1-572 <MIN>  
A/Cross-references: UNIPROT:P47942; GB:246882; NID:G599965; PIDN:CAA86981.1; PID:G599966  
A/Experimental source: strain Sprague-Dawley; clone TOAD64 (turned on after division, 64)  
A/Note: submitted to the EMBL Data Library, November 1994  
A/Note: in Genbank entry R70AD64, release 113.0, the source is designated as Rattus rat  
C/Superfamily: allantoinase; Bacillus dihydroorotase homology

Query Match	78.8%	Score 2349;	DB 2;	Length 572;
Best Local Similarity	76.0%;	Pred. No. 5.3e-163;		
Matches 434;	Conservative	70;	Mismatches 67;	Indels 0;

Qy	1	MSFGKKSIPRIITSDRLIRGGIRVNDQSFYADVHVVDGLIKQIGENLIVPGGIKTIDA	60
Db	1	MSYQKKNIPIRITSDDLIRGGIRVNDQSFYADIVMEDGLIKQIGENLIVPGGVKTIEA	60
Qy	61	HGLMVLPGVDVHTRLOMPVLGTPADDPCQGTAKAALAGGTTMLDHHVFPDTCVSLAAY	120
Db	61	HSRMLIPGGIDVHTRFQMPQDQMTSADDFPQGTAKAALAGGTTMLDHHVPEPQTSLLAAP	120
Qy	121	EQWRERADSAACDYSILVDITRWHSIKLEALVKEKGVNSFLVPMAYKDRQCQSDSQ	180
Db	121	DQWRWADSKSCDYSILVDITRWHSIQEALVKEALVKEGVNSFLVPMAYKDRQFQLTDSQ	180
Qy	181	MYEIFSIRDLGALAQVHAENGDIVESQKRLLELGTGTPGCVHLSHPVEVEAEAVRAY	240
Db	181	IYEVLSVIRDIGALAQVHAENGDIIEBEQRIILDGITGPEGHVLSRPEVEAEAVNRSI	240
Qy	241	TIQAQNCPLVYTKVMSKGAADAIQAQRGVVVFGEPIITASIGTDSHWSKNWAKAAA	300
Db	241	TIANTQNCPLVYTKVMSKSAEAVIAQARKKGTVVYGEPIITASIGTDSHWSKNWAKAAA	300
Qy	301	FVTSPPVNPDPPTADHLTCLLSSGLQVTSAGHCTFTTAQKAVGKONFALIPEGTNGIEE	360
Db	301	FVTSPPVNPDPPTDFNLSSGLQVTSAGHCTFTTAQKAVGKONFTLIPGTNGITEE	360
Qy	361	RMSVMWKECVASGQNDENEFVAVTSTNAAKIFNYPYRKGKRVAVGSDADLVINPKATKII	420
Db	361	RMSVIMDKAVVTKGMDENQFVAVTSTNAAKVFNLYPRKGRISVSGSDADLVIMDPDSVKTI	420
Qy	421	SAKTHNLNVEYNIPFEGVECEGAPAVVISQGRVALEDGQMFVTPGAGRFVPRKTPPDPVYK	480
Db	421	SAKTHNSALEYNIPFEGMECEGRSPLVVISQGRVILEDTGLHVTGEGSGRYIIPKQFPDPVYK	480
Qy	481	RIKARNRLAEIHGVPRLGYDGPVHEVMVPKPGSGAPARASCFOKISVPPVRNLHQSGFS	540
Db	481	RIKARSRLAELRGVPRGLYDGPVCEVSVTKVTTPASSAKTSPAKQQAQPPVRNLHQSGFS	540
Qy	541	LSGSQADDDHIAARRTAQKIMAPPGRSNTSL	571



Db 540 LSGTQVDEGV-RSASKRIIVAPGGRSNTLSL 570

RESULT 6  
JCS316  
dhHydroxyrimidinase related protein 1 - human  
C/Date: 01-May-1997 #sequence\_revision 18-Jul-1997 #text\_change 09-Jul-2004  
A/Accession: JCS316, S58890  
R/Hamajima, N.; Matsuda, K.; Sakata, S.; Tamaki, N.; Sasaki, M.; Nonaka, M.  
Gene 180, 157-163, 1996  
A/Title: A novel gene family defined by human dihydroxyrimidinase and three related proteins  
A/Reference number: JCS315; MUID:97128821; PMID:8973361  
A/Accession: JCS316  
A/Status: nucleic acid sequence not shown  
A/Molecule type: mRNA  
A/Residues: 1-572 <HAM>  
A/Cross-references: UNIPROT:Q14194; DDBJ:D78012; MID:g1330237; PIDN:BAAL1190.1; PID:g1330237  
A/Experimental source: fetal brain  
R/Goshima, Y.; Nakamura, F.; Strittmatter, P.; Strittmatter, S.M.  
Nature 376, 509-514, 1995  
A/Title: Collapsin-induced growth cone collapse mediated by an intracellular protein related to semaphorin 3A  
A/Reference number: S5889; MUID:95364923; PMID:7637782  
A/Accession: S58890  
A/Status: preliminary; nucleic acid sequence not shown; translation not shown  
A/Molecule type: mRNA  
A/Residues: 64-503, 'H', 505-572 <GOS>  
A/Cross-references: EMBL:U17278; MID:9882148; PIDN:AAA93201.1; PID:9882149  
A/Note: The nucleotide sequence was submitted to the EMBL Data Library, November 1994  
C/Comment: This protein is involved in nervous system development.  
C/Genetics:  
A/Gene: GDB:CRMPI; DRP-1; DPVSL1  
A/Cross-references: GDB:S585714  
C/Superfamily: allantoinase; Bacillus dihydroorotase homology

Query Match 72.3%; Score 2157; DB 2; Length 572;  
Best Local Similarity 69.2%; Pred. No. 5e-149;  
Matches 395; Conservative 85; Mismatches 91; Indels 0; Gaps 0;

Qy 1 MSFGKGSIPRISDRLLIRGGIRVNDQSFYADVHVDEGLIKQIGENLVPGIKTIDA 60  
Db 1 MSYQKKSPHITSRLIRGGIRVNDQSFYADVHVDEGLIKQIGENLVPGIKTIDA 60

Qy 61 HGLMVLPGGVVHTRLOMPLVGTADDPCQGTAAALAGTMTILDHVPDTCVSLAAY 120  
Db 61 NGRMVLPGGVVHTRLOMPLVGTADDPCQGTAAALAGTMTILDHVPDTCVSLAAY 120

Qy 121 EWRERADSAACDYSILHVDITWHSIIEEALVKEKGVNSFLVFMAYKQRCQSDSQ 180  
Db 121 EKWHEADTKSCDYSILHVDITWSDGVRELEVLVQDKGVNSFQVYMAKQVQMSDSQ 180

Qy 181 MYEIFSIRDLGALAOVHAENGDI VEEOKRLLELGTGPBGHVLSPHREVEAEVYRAV 240  
Db 181 LYEAFTFLGLGAVILVHAENGDLIAEQKRILEMGITGPEGHALSRPELEAEVRAI 240

Qy 241 TIAQANCPVYTKVMSKGAADIAQAKRGVVVFGPEPTASLGTDSHYMSKNWAKAA 300  
Db 241 TIAGRINCPVYTKVMSKGAADIAALARKKGPLVFGPEPTAASLGTDSHYMSKNWAKAA 300

Qy 301 FVTSPPVNPDPPTADHLTLCLSSGDIQVTSAGHCTFTTAQAVGKNFALIPGTNGIEE 360  
Db 301 FVTSPPVNPDPPTADHLTLCLSSGDIQVTSAGHCTFTTAQAVGKNFALIPGTNGIEE 360

Qy 361 RMSMWKCVASGMDENEFVATSTNAKIFNFPYPRKGVAVGSDADLVINPNKATKII 420  
Db 361 RMTVMWKAATGMDENEFVATSTNAKIFNLYPRKGVAVGSDADLVINPNKATKII 420

Qy 421 SATHNLNVEYNIPFEGVECRGAPAVTSQGRVALEDGKMPVTPGAGRFVPRKTFPDPVYK 480  
Db 421 TAKSHKSAVEYNIPFEGVECRGAPAVTSQGRVALEDGKMPVTPGAGRFVPRKTFPDPVYK 480

Qy 481 RIKARNRLABIHGVPRGLYDGPVHEVMVPAKPGSGAPARASCPCGKISVPPVRLHQSGFS 540

Db 481 RVKIRNKVFLQGVSRGMVDGVPYVPATPKYATPAPSAKSPSKHQPPPIRLHQSNFS 540

Qy 541 LSGSQADDDHIAARTAAQKIMAPPGRSNTLSL 571  
Db 541 LSGAQIDDDNPRRTGHRIVAPPGRSNTLSL 571

RESULT 7  
JCS315  
dhHydroxyrimidinase (EC 3.5.2.2) - human  
N/Alternate names: 5,6-dihydroxyrimidine amidohydrolase; Hydantoinase  
C/Species: Homo sapiens (man)  
C/Date: 01-May-1997 #sequence\_revision 18-Jul-1997 #text\_change 09-Jul-2004  
A/Accession: JCS315  
R/Hamajima, N.; Matsuda, K.; Sakata, S.; Tamaki, N.; Sasaki, M.; Nonaka, M.  
Gene 180, 157-163, 1996  
A/Title: A novel gene family defined by human dihydroxyrimidinase and three related proteins  
A/Reference number: JCS315; MUID:97128821; PMID:8973361  
A/Accession: JCS315  
A/Status: nucleic acid sequence not shown  
A/Molecule type: mRNA  
A/Residues: 1-519 <HAM>  
A/Cross-references: UNIPROT:Q14117; DDBJ:D78011; MID:g22339965; PIDN:BAAL1189.1; PID:g1330237  
A/Experimental source: liver  
C/Comment: This enzyme is Zn2+-metalloenzyme, and the second enzyme involved in uracil aryl  
dithymine to N-carbamyl-beta-aminoisobutyrate. It catalyzes the hydrolysis of a variety  
C/Genetics:  
A/Gene: GDB:DPYS; DHPase  
A/Cross-references: GDB:S5885803  
C/Complex: homotetramer  
C/Superfamily: allantoinase; Bacillus dihydroorotase homology  
C/Keywords: hydrolase

Query Match 52.2%; Score 1556; DB 2; Length 519;  
Best Local Similarity 59.3%; Pred. No. 2.4e-105;  
Matches 292; Conservative 70; Mismatches 126; Indels 4; Gaps 1;

Qy 16 RLLIIGRRVNDQSFYADVHVDEGLIKQIGENLVPG----GIKTIDAHGLMVLPGGV 71  
Db 6 RLLIIGRRVNDQSFYADVHVDEGLIKQIGENLVPGGVAGAPAGRLVLDAGKLVLPGGID 65

Qy 72 VHTRLQMPVLGTMPTADDPCQGTAAALAGTMTILDHVPDTCVSLAAYEOWREADSA 131  
Db 66 THTHQFPFMGSRSDDDFHQGTAAALSGGTTMIDPAIPQKGSLEAEFTWRSWADPKV 125

Qy 132 CDYSILHVDITWHSIIEEALVKEKGVNSFLVFMAYKQRCQSDSQSYEIFSIRDL 191  
Db 126 CDYSILHVAVTWMSDQVKEEMKILVQDKGVNSFKMFAYKOLYMTDLELYEAFSRCKE 185

Qy 192 GALAOVHAENGDI VEEOKRLLELGTGPBGHVLSPHREVEAEVYRAVTTIAQANCP 251  
Db 186 GAIQVHAENGDLIAAGAKKMLALGITGPEGHLCRPEAEAEATLITIASAVNCP 245

Qy 252 VTKVMSKGAADIAQAKRGVVVFGPEPTASLGTDSHYMSKNWAKAAAFVTSPPVNP 311  
Db 246 IVTVMSKSAKVIADARRDKVYGEPTAASLGTDSHYMSKNWAKAAAFVTSPPVNP 305

Qy 312 TTADHLTLCLSSGDIQVTSAGHCTFTTAQAVGKNFALIPGTNGIEERMSWWEKCV 371  
Db 306 STPDFLMLNLLANDLTLTGTNDCTENTCOKALGKDDFTKIPNGVNGVEDRMSVWEKGV 365

Qy 372 SGMDENEFVATSTNAKIFNFPYPRKGVAVGSDADLVINPNKATKIIISAKTHNLNVE 431  
Db 366 SGMDENEFVATSTNAKIFNLYPRKGVAVGSDADLVINPNKATKIIISAKTHNLNVE 425

Qy 432 NTFEGVECRGAPAVTSQGRVALEDGKMPVTPGAGRFVPRKTFPDPVYKRIKARNRL 491  
Db 426 NTFEGVECRGAPAVTSQGRVALEDGKMPVTPGAGRFVPRKTFPDPVYKRIKARNRL 485

Qy 492 HGVPRGLYDGPV 503  
Db 486 TPVERAPYKGEV 497



RESULT 8  
S70581  
dihydroxyrimidine - rat  
C;Species: Rattus norvegicus (Norway rat)  
C;Date: 14-Feb-1997 #sequence\_revision 13-Mar-1997 #text\_change 09-Jul-2004  
C;Accession: S70581  
R;Matsuda, K.; Sakata, S.; Kaneko, M.; Hamajima, N.; Nonaka, M.; Sasaki, M.; Tamaki, N.  
Biochim. Biophys. Acta 1307, 140-144, 1996  
A;Title: Molecular cloning and sequencing of a cDNA encoding dihydroxyrimidine from the  
A;Reference number: S70581; MUID:56283806; PMID:8679696  
A;Accession: S70581  
A;Status: preliminary  
A;Molecule type: mRNA  
A;Residues: 1-519 <MAT>  
A;Cross-references: UNIPROT:Q63150; EMBL:D63704; NID:g1378018; PIDN:BAA09833.1; PID:d101  
C;Superfamily: allantoinase; Bacillus dihydroorotase homology

Query Match 52.0%; Score 1550; DB 2; Length 519;  
Best Local Similarity 58.4%; Pred. No. 6.6e-105;  
Matches 293; Conservative 69; Mismatches 132; Indels 8; Gaps 2;

QY 15 DRLIRGRIVNDQSPYADVHVEDGLIKOIGENLIVPG---GIKTIDAHGLMVLPGGV 70  
DB 5 ERLIRGRVNDQSPYADVHVEDGLIKOIGENLIVPG---GIKTIDAHGLMVLPGGI 64  
QY 71 DVHTRIQMPVLGMPADDFCQGTAAALAGGTTMILDHVPFDTGVSLLAAEYQWRERADSA 130  
DB 65 DTHTHMQPMPGSGVDDFHQGTAAALAGGTTMIDFAIPQKSSLIIEAFETWRNWDPK 124  
QY 131 ACCDYSILHVDITWHESIKKEELALVKEKGVNSFLPMAYKDRQCSDSOMYEIFSIIRD 190  
DB 125 VCCDYSILHVAVTWSDKVEEMKTLAQDKGVNSFKMFMYKOLYMQDQOMYAAFSQCKE 184  
QY 191 LGAALQVHAENGDI VEEEOKRLLELGTGPGEHVLSPHEVEAEAVYRAVTIAQANCP 250  
DB 185 IGAIQVHAENGDLIAEGAKKMLALGTGPGEHLCRPEAEAEATLRAITIASAVNCPL 244  
QY 251 YVTKVMSKGAADAIAQAKRGVVVFGPITASLGTGSHYMSKNWAKAAAFVTSPPVNP 310  
DB 245 YIVHVMKSNAKVIADAKRGKVYGEPIAAGLTGTGYWKEWRAHAAHVWGPIRLPD 304  
QY 311 PTTADHLTCLSSGDLQVTSAGHCTFTTAQKAVGKNFALIPGTNGIEERMSMWKCV 370  
DB 305 PSTPGFLNLLANGDLTTTGSDNCTNTCQALGKDDFTKIPNGVNGVEDRMSVWEKGV 364  
QY 371 ASGKMDENRPAVTSNAAKIFNYPKRGVAVGSDADLVINPKATKIISAKTHNLVE 430  
DB 365 HSGKMDENRPAVTSNAAKIFNLYPKGRIAVGSDADLVINPKATKIISAKTHNLVE 424  
QY 431 YNIFEGVECRGAPAVTSQGRVALEDGKMFVTPGAGRFVPRKTPDFVYKRIKARNLAE 490  
DB 425 FNIFEGMVCHGVPLVTSIRGTVYVYAGVPTVAGHGKFIPTQPPAFIYKRVKQDQCT 484  
QY 491 IHGVPRLGYDGPVHEVMVPAKP 512  
DB 485 IIPVKRAPHYKGEV---ITLKP 502

RESULT 9  
T20007  
hypothetical protein C47E12.8 - Caenorhabditis elegans  
C;Species: Caenorhabditis elegans  
C;Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 09-Jul-2004  
C;Accession: T20007  
R;Coles, L.  
submitted to the EMBL Data Library, January 1996  
A;Reference number: Z19210  
A;Accession: T20007  
A;Status: preliminary; translated from GB/EMBL/DBJ  
A;Molecule type: DNA  
A;Residues: 1-520 <WIL>  
A;Cross-references: UNIPROT:Q18677; EMBL:Z68882; PIDN:CAA93104.1; GSPDB:GN00022; CESP:C4

A;Experimental source: clone C47E12  
C;Genetics:  
A;Gene: CESP:C47E12.8  
A;Map position: 4  
A;Introns: 29/2; 89/3; 147/1; 232/3; 310/2; 414/3  
C;Superfamily: allantoinase; Bacillus dihydroorotase homology

Query Match 41.8%; Score 1246; DB 2; Length 520;  
Best Local Similarity 45.4%; Pred. No. 8.9e-83;  
Matches 251; Conservative 90; Mismatches 174; Indels 38; Gaps 3;

QY 17 LLIRGRIVNDQSPYADVHVEDGLIKOIGENLIVPGGKITIDAHGLMVLPGGVDTHTRL 76  
DB 3 LLIRGRIVNDQSPYADVHVEDGLIKOIGENLIVPGGKITIDAHGLMVLPGGIDPHTHM 62  
QY 77 QNPVLGMPADDFCQGTAAALAGGTTMILDHVPFDTGVSLLAAEYQWRERADSAACDYS 136  
DB 63 QNPVLGMPADDFCQGTAAALAGGTTMILDHVPFDTGVSLLAAEYQWRERADSAACDYS 122  
QY 137 LHVDITRWHESIKKEELALV-KEKGVNSFLPMAYKDRQCSDSOMYEIFSIIRDGLA 195  
DB 123 LSWAITSWGPETAKEMEIVTGAEGYINSFKFLAYAGVFMVRDEEFYQGMIOCAKLALA 182  
QY 196 QVHAENGDIVREEOKRLLELGTGPGEHVLSPHEVEAEAVYRAVTIAQANCPYTKV 255  
DB 183 RVHAENGSVIARCEHLLSSGITGPGEHTQSRPELEAEATFRACMASQANCPYVHV 242  
QY 256 MSKGAADAIAQAKRGVVVFGPITASLGTGSHYMSKNWAKAAAFVTSPPVNPDP 315  
DB 243 MSKGAADAIAQAKRGVVVFGPITASLGTGSHYMSKNWAKAAAFVTSPPVNPDP 302  
QY 316 HTCLSSGDLQVTSAGHCTFTTAQKAVGKNFALIPGTNGIEERMSMWKCVASGKM 375  
DB 303 ALMKLLAAGELHLTATDNCTFDCQKSLGKDDFTKIPNGVNGVEDRMSVWMDKGVHAGII 362  
QY 376 DENEFAVTVSTNAAKIFNYPKRGVAVGSDADLVINPKATKIISAKTHNLVEINFI 435  
DB 363 DPMRFVAVTVSTNAAKIFNYPKRGVAVGSDADLVINPKATKIISAKTHNLVEINFI 422  
QY 436 GVECRGAPAVTSQGRVALEDGKMFVTPGAGRFVPRKTPDFVYKRIKARNLAEIHGVP 495  
DB 423 GNVHGVCHGVPLVTSIRGTVYVYAGVPTVAGHGKFIPTQPPAFIYKRVKQDQCT 477  
QY 496 RGLYDGPVHEVMVPAKP 512  
DB 478 -----PVKIDRIPIYEPS-----ALQTPDANANIVVKAP 505  
QY 556 QKIMAPPGGRSNI 568  
DB 506 VRAAIPPGGASSI 518

RESULT 10  
T23968  
hypothetical protein R06C7.3 - Caenorhabditis elegans  
C;Species: Caenorhabditis elegans  
C;Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 02-Sep-2000  
C;Accession: T23968  
R;Gardner, A.  
submitted to the EMBL Data Library, April 1996  
A;Reference number: Z19825  
A;Accession: T23968  
A;Status: preliminary; translated from GB/EMBL/DBJ  
A;Molecule type: DNA  
A;Residues: 1-544 <WIL>  
A;Cross-references: EMBL:Z71266; PIDN:CAA95842.1; GSPDB:GN00019; CESP:R06C7.3  
C;Genetics:  
A;Gene: CESP:R06C7.3  
A;Map position: 1  
A;Introns: 103/3; 135/2; 195/3; 255/1; 325/3; 418/2; 440/2; 522/3  
C;Superfamily: allantoinase; Bacillus dihydroorotase homology





A/Accession: PC2206  
A/Molecule type: protein  
A/Residues: 1-20 <MU2>  
C/Comment: This enzyme is a metalloenzyme and the oligomeric structure is required for a  
C/Superfamily: allantoicase; Bacillus dihydroorotase homology  
C/Keywords: hydrolase

Query Match 28.1%; Score 838; DB 1; Length 471;  
Best Local Similarity 40.0%; Pred. No. 3.8e-53;  
Matches 183; Conservative 80; Mismatches 191; Indels 4; Gaps 4;

QY	18	LIRGRIVNDQSFYADVHVEDLIQIGENLIVPGGIKTIDAHGLMLVPGGVVDVHTRLQ	77
DB	4	LIRGRIVNDQSFYADVHVEDLIQIGENLIVPGGIKTIDAHGLMLVPGGVVDVHTRLQ	77
QY	78	MPVLGMPADDDFCQGGTKAALAGTTMLDHFVDPDGVSLAAYEQWRERADSAACCDYSL	137
DB	63	MPFGGTVTKDDPESGTTAAAFGGTTIIDFCLTNKGEPLKKAETWHNKATGKAVIDYGF	122
QY	138	HVDITRWHSIKKELEALVKEKGVNSPLVPMAYKDRQCQSDSOMYEIPIIIRDLGALAOV	197
DB	123	HLWISITDDVLEELPKVIEEGITSFVKVPMAYKQVFOADDGTLXETLVAAKELGALVMV	182
QY	198	HAENGDIVBEEQKRLLELGTGPEGHVLSHPPEVEAEAVVRAVTIAKQANCPLYVTKVMS	257
DB	183	HAENGDIVDLYTKALEDSHTDPIYHALTRPPELEGEATGRACQLTELASQLYVVHVSC	242
QY	258	KGAADIAQAQRGVVVFGEPTASLGTGSHYMSKNMAKAAAFVTSPPVNPDPPTTADHL	317
DB	243	AQAVEKIAEARNKGLNMGETCPQLVLDQSYLEKDNF-EGAKYVWSPPLR-EKWHQEV	300
QY	318	TCLLSGDLQVTSAGHCTFT-TAQKAVGKDNFALIPGTNGIERMSMWKCVASGKMD	376
DB	301	WNAKNGQLQLGSDQCSDFKQKELGRGDFTKIPNGGPIEDRVSIILFSEGVKKGRIT	360
QY	377	ENEFVAVTSTNAKIFNFPYPRGRVAVGSDADLVINPKATKIIISAKTNLNVENIFEG	436
DB	361	LNQFVDIVSTRIAKLFLPKKGTIAVGADADLVIFDPTVERVISAEETHHMAVDYNPFEG	420
QY	437	VECRGAPAVVIGGRVALSDGKMFVTPGAGRFVPRKTF	474
DB	421	MKVTGEPVSLCRGEPVVRDKQFVGKPGYQYVVKRAKY	458

Search completed: September 24, 2005, 17:39:34  
Job time : 274 secs



CC human immune deficiency viruses. CRMP is also used to identify agents  
 CC potentially useful for treating prion diseases and to detect abnormal  
 CC expression or localization of CRMP in immune system cells, for diagnosis  
 CC and prognosis of diseases of the immune system. ABP97907-11 represent  
 CC human CRMP proteins  
 XX  
 SQ Sequence 572 AA;

Query Match 99.9%; Score 2978; DB 6; Length 572;  
 Best Local Similarity 99.8%; Pred. No. 5.4e-277;  
 Matches 571; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSFQGGKSIPTISDRLLIRGGRIVNDQSFYADVHVEDGLIKQIGENLIVPGIKTIDA 60  
 DB 1 MSFQGGKSIPTISDRLLIRGGRIVNDQSFYADVHVEDGLIKQIGENLIVPGIKTIDA 60  
 QY 61 HGLMWLPGGVDVHTRIQMPVLGWTPTADDFCQGTAAALAGTTMILDHVPDPTGVSLAAY 120  
 DB 61 HGLMWLPGGVDVHTRIQMPVLGWTPTADDFCQGTAAALAGTTMILDHVPDPTGVSLAAY 120  
 QY 121 EWRERADSAACCDYSLHVDITRWHSIKKELEALVKEGVNSFLVFMAYKORCQSDSQ 180  
 DB 121 EWRERADSAACCDYSLHVDITRWHSIKKELEALVKEGVNSFLVFMAYKORCQSDSQ 180  
 QY 181 MYEIPSIIRDGLAQAQVHAENGDIVEEROKRLELIGITGPEGHVLSPBEVEAAVYRAV 240  
 DB 181 MYEIPSIIRDGLAQAQVHAENGDIVEEROKRLELIGITGPEGHVLSPBEVEAAVYRAV 240  
 QY 241 TIAKQANCPLVYTKVMSKGAADAIQAQKRGVVVFGEPIITASIGTDSHYWSKNWAKAA 300  
 DB 241 TIAKQANCPLVYTKVMSKGAADAIQAQKRGVVVFGEPIITASIGTDSHYWSKNWAKAA 300  
 QY 301 FVTSPPVNPDPPTADHLTCLLSSGDIQVTSAGHCTFTTAQKAVGKONFALIPGTNGIEE 360  
 DB 301 FVTSPPVNPDPPTADHLTCLLSSGDIQVTSAGHCTFTTAQKAVGKONFALIPGTNGIEE 360  
 QY 361 RMSWWEKCVASGMDENEFVATSTNAKIFNYPYRKGKRVAVGSDADLVINWPKATKII 420  
 DB 361 RMSWWEKCVASGMDENEFVATSTNAKIFNYPYRKGKRVAVGSDADLVINWPKATKII 420  
 QY 421 SAKTHNLNVEYNIFEGVECRGAPAVVISQGVVALEDKGMFVTPGAGRFVPRKTFPDPVYK 480  
 DB 421 SAKTHNLNVEYNIFEGVECRGAPAVVISQGVVALEDKGMFVTPGAGRFVPRKTFPDPVYK 480  
 QY 481 RIKARNRLAEIHGVPRGLYDGPVHEVMVPAKPGSGAPARASCPSGKISVPPVRNLHQSGFS 540  
 DB 481 RIKARNRLAEIHGVPRGLYDGPVHEVMVPAKPGSGAPARASCPSGKISVPPVRNLHQSGFS 540  
 QY 541 LSGSQADDDHIARTAKINAPPGGRSNTLS 572  
 DB 541 LSGSQADDDHIARTAKINAPPGGRSNTLS 572

## RESULT 2

ID ADP65307  
 AC ADP65307 standard; protein; 572 AA.  
 AC ADP65307;  
 DT 12-AUG-2004 (first entry)  
 XX Human dihydropyrimidinase-like 4.  
 XX autoimmune disease; arthritis; gene expression analysis;  
 KW rheumatoid arthritis; collagen-induced; immunosuppressive; anti-rheumatic;  
 KW antiarthritis; osteopathic; antigout; anti-inflammatory; dermatological;  
 KW immunomodulatory; lupus; ankylosing spondylitis; fibrositis;  
 KW fibromyalgia; osteoarthritis; gout; juvenile rheumatoid arthritis;  
 KW immune; human.  
 OS Homo sapiens.  
 OS  
 XX WO2003072827-A1.  
 PN

XX 04-SEP-2003.  
 PD  
 XX 31-OCT-2002; 2002WO-US035433.  
 PF  
 XX 31-OCT-2001; 2001US-0336220P.  
 PR  
 XX (CHIL-) CHILDREN'S HOSPITAL MEDICAL CENT.  
 PA  
 XX Hirsch R, Thornton SL;  
 PI  
 XX WPI: 2003-712740/67.  
 DR  
 XX GENBANK; NP\_006417.  
 XX  
 DR  
 XX  
 XX  
 PT Diagnosing and analyzing autoimmune disease using gene expression  
 PT profiles and microarray technology, useful for diagnosing and treating  
 PT rheumatoid arthritis, lupus, fibrositis, osteoarthritis, fibromyalgia and  
 PT gout.  
 PT  
 XX  
 PS Disclosure; Page; 56pp; English.

The invention relates to a novel method for diagnosing and analyzing  
 CC autoimmune disease or arthritides. The method comprises obtaining a  
 CC patient sample containing mRNA, analyzing gene expression using the mRNA  
 CC that results in a gene expression signature of the mRNA, and using that  
 CC gene expression signature to diagnose or analyze the autoimmune disease  
 CC or arthritides in the patient, where gene expression of at least 60% of  
 CC the genes correlates with that of the gene signature. The invention  
 CC further comprises: a treatment of rheumatoid arthritis; identification of  
 CC genes for targeting in the treatment of rheumatoid arthritis in a mammal  
 CC other than a mouse; diagnosis of rheumatoid arthritis in a mammal; an  
 CC array or gene chip, specific for rheumatoid arthritis; diagnosis or  
 CC analyses of autoimmune disease or rheumatoid arthritis; screening the  
 CC efficacy of a candidate drug in vitro for the treatment of collagen-  
 CC induced arthritis; and reducing the symptoms associated with collagen-  
 CC induced arthritis. The compositions of the invention have the following  
 CC activities: immunosuppressive, anti-rheumatic, antiarthritis, osteopathic,  
 CC antigout, anti-inflammatory, dermatological, and immunomodulatory. The  
 CC methods and compositions of the present invention are useful for  
 CC diagnosing and treating autoimmune disease or arthritides, such as  
 CC rheumatoid arthritis, lupus, ankylosing spondylitis, fibrositis,  
 CC fibromyalgia, osteoarthritis, gout, juvenile rheumatoid arthritis, and an  
 CC immune disease caused by an infectious agent. This sequence represents a  
 CC protein sequence relating to the genes used in the analysis and treatment  
 CC of autoimmune diseases or arthritides. Note: This sequence is not shown  
 CC in the specification. It has been supplied in an electronic format from  
 CC WIPO.

XX Sequence 572 AA;

Query Match 99.9%; Score 2978; DB 7; Length 572;  
 Best Local Similarity 99.8%; Pred. No. 5.4e-277;  
 Matches 571; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSFQGGKSIPTISDRLLIRGGRIVNDQSFYADVHVEDGLIKQIGENLIVPGIKTIDA 60  
 DB 1 MSFQGGKSIPTISDRLLIRGGRIVNDQSFYADVHVEDGLIKQIGENLIVPGIKTIDA 60  
 QY 61 HGLMWLPGGVDVHTRIQMPVLGWTPTADDFCQGTAAALAGTTMILDHVPDPTGVSLAAY 120  
 DB 61 HGLMWLPGGVDVHTRIQMPVLGWTPTADDFCQGTAAALAGTTMILDHVPDPTGVSLAAY 120  
 QY 121 EWRERADSAACCDYSLHVDITRWHSIKKELEALVKEGVNSFLVFMAYKORCQSDSQ 180  
 DB 121 EWRERADSAACCDYSLHVDITRWHSIKKELEALVKEGVNSFLVFMAYKORCQSDSQ 180  
 QY 181 MYEIPSIIRDGLAQAQVHAENGDIVEEROKRLELIGITGPEGHVLSPBEVEAAVYRAV 240  
 DB 181 MYEIPSIIRDGLAQAQVHAENGDIVEEROKRLELIGITGPEGHVLSPBEVEAAVYRAV 240  
 QY 241 TIAKQANCPLVYTKVMSKGAADAIQAQKRGVVVFGEPIITASIGTDSHYWSKNWAKAA 300  
 DB 241 TIAKQANCPLVYTKVMSKGAADAIQAQKRGVVVFGEPIITASIGTDSHYWSKNWAKAA 300

QY 301 FVTSPPVNPDPPTADHLTCLSSGDLQVTSAGHCTTTAAKAVGKDNFALIPGTINGIEE 360  
DB 301 FVTSPPVNPDPPTADHLTCLSSGDLQVTSAGHCTTTAAKAVGKDNFALIPGTINGIEE 360  
QY 361 RMSMWKCVASGKMDNEFVATSTNAAKIFNFPYPRKGRVAVGSDADLVINPKATKII 420  
DB 361 RMSMWKCVASGKMDNEFVATSTNAAKIFNFPYPRKGRVAVGSDADLVINPKATKII 420  
QY 421 SAKTHNLNVEYNIPEGVECRGAPAVVISQGRVALEDGKMPVTPGACGRFVPRKTFPDPVYK 480  
DB 421 SAKTHNLNVEYNIPEGVECRGAPAVVISQGRVALEDGKMPVTPGACGRFVPRKTFPDPVYK 480  
QY 481 RIKARNRLAEIHGVPRLGYPVHVMVPAKPGSGAPARASCPCGKISVPPVRLHQSGFS 540  
DB 481 RIKARNRLAEIHGVPRLGYPVHVMVPAKPGSGAPARASCPCGKISVPPVRLHQSGFS 540  
QY 541 LSGSQADDDHARRTAQKIMAPPGGRSNTLS 572  
DB 541 LSGSQADDDHARRTAQKIMAPPGGRSNTLS 572

## RESULT 3

ADQ18668 standard; protein; 572 AA.

XX AC ADQ18668;  
XX DT 26-AUG-2004 (first entry)  
XX DE Human soft tissue sarcoma-upregulated protein - SEQ ID 1487.  
XX KW soft tissue sarcoma; cytostatic; gene therapy; vaccine; screening; human.  
XX OS Homo sapiens.  
XX FN WO2004048938-A2.  
XX PD 10-JUN-2004.  
XX PF 26-NOV-2003; 2003WO-US038193.  
XX PR 26-NOV-2002; 2002US-0429739P.  
XX PA (PROT-) PROTEIN DESIGN LABS INC.  
XX PI Aziz N, Ginsburg WM, Zlotnik A;  
XX DR WPI; 2004-441208/41.

XX PT Early detection of soft tissue sarcoma comprises determining expression  
XX PT of a gene in a first soft tissue sample and a normal soft tissue sample  
XX PT and comparing the gene expression, also useful in treating soft tissue  
XX PT sarcoma;

XX PS Example 2; SEQ ID NO 1487; 210pp; English.

XX CC The invention relates to a novel method for detecting soft tissue sarcoma  
XX CC which comprises obtaining a first soft tissue sample from an individual  
XX CC and a normal soft tissue sample from the same or different individual,  
XX CC determining the expression of a gene in both samples and comparing the  
XX CC expression of the gene in both soft tissue samples, where a higher level  
XX CC of protein expression in the first soft tissue sample indicates the  
XX CC presence of soft tissue sarcoma. The method of the invention has  
XX CC cytostatic applications and may be useful for detecting soft tissue  
XX CC sarcoma, possibly via gene therapy or vaccine production. The nucleic  
XX CC acid sequences may be useful in diagnostic and screening applications.  
XX CC The current sequence is that of a human soft tissue sarcoma-upregulated  
XX CC protein of the invention. The current sequence is not shown within the  
XX CC specification per se but was submitted in CD format by the inventor.

XX SQ Sequence 572 AA;

Query Match 99.9%; Score 2978; DB 8; Length 572;  
Best Local Similarity 99.8%; Pred. No. 5.4e-277;  
Matches 571; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MSFQGGKSPRTSDRLIRGGRIVNDQSFYADVHVHEDGLIKQIGENLIVPGIKTIDA 60  
DB 1 MSFQGGKSPRTSDRLIRGGRIVNDQSFYADVHVHEDGLIKQIGENLIVPGIKTIDA 60  
QY 61 HGLMWLPGVVDVHTRQLQMPVLGWTTPADDFCOQTQAALAGGTTMLDHPVDPDTCVSLAAY 120  
DB 61 HGLMWLPGVVDVHTRQLQMPVLGWTTPADDFCOQTQAALAGGTTMLDHPVDPDTCVSLAAY 120  
QY 121 EQRERADSAACDYSLSHVDITRWHSIKEEALVKKGVSFLVFMAYKDCQCSDSQ 180  
DB 121 EQRERADSAACDYSLSHVDITRWHSIKEEALVKKGVSFLVFMAYKDCQCSDSQ 180  
QY 181 MYEISIIIRDLGALAQVHAENGDIVEEEOKRLLLELGTGPEGHVLSHPBEVEAEVYRAV 240  
DB 181 MYEISIIIRDLGALAQVHAENGDIVEEEOKRLLLELGTGPEGHVLSHPBEVEAEVYRAV 240  
QY 241 TIAQANCPLYTVTKMSKGAADAIQAQRGVVVFGEPIITASLGTDGSHYWSKNWAKAA 300  
DB 241 TIAQANCPLYTVTKMSKGAADAIQAQRGVVVFGEPIITASLGTDGSHYWSKNWAKAA 300  
QY 301 FVTSPPVNPDPPTADHLTCLSSGDLQVTSAGHCTTTAAKAVGKDNFALIPGTINGIEE 360  
DB 301 FVTSPPVNPDPPTADHLTCLSSGDLQVTSAGHCTTTAAKAVGKDNFALIPGTINGIEE 360  
QY 361 RMSMWKCVASGKMDNEFVATSTNAAKIFNFPYPRKGRVAVGSDADLVINPKATKII 420  
DB 361 RMSMWKCVASGKMDNEFVATSTNAAKIFNFPYPRKGRVAVGSDADLVINPKATKII 420  
QY 421 SAKTHNLNVEYNIPEGVECRGAPAVVISQGRVALEDGKMPVTPGACGRFVPRKTFPDPVYK 480  
DB 421 SAKTHNLNVEYNIPEGVECRGAPAVVISQGRVALEDGKMPVTPGACGRFVPRKTFPDPVYK 480  
QY 481 RIKARNRLAEIHGVPRLGYPVHVMVPAKPGSGAPARASCPCGKISVPPVRLHQSGFS 540  
DB 481 RIKARNRLAEIHGVPRLGYPVHVMVPAKPGSGAPARASCPCGKISVPPVRLHQSGFS 540  
QY 541 LSGSQADDDHARRTAQKIMAPPGGRSNTLS 572  
DB 541 LSGSQADDDHARRTAQKIMAPPGGRSNTLS 572

## RESULT 4

ABG32231 standard; protein; 572 AA.

XX AC ABG32231;  
XX DT 05-NOV-2002 (first entry)  
XX DE Human Ulip4/CRMP3 protein.  
XX KW Human; Ulip4; CRMP3, collapsin response mediator protein;  
XX KW Unc-33-like protein; neurodegenerative disease; Alzheimer's disease;  
XX KW paraneoplastic neurodegenerative disease; PND; myelination;  
XX KW demyelination; remyelination; myelin disorder; multiple sclerosis;  
XX KW autoimmune neurodegenerative disorder; HTLV-1 associated myelopathy;  
XX KW human T lymphocyte virus 1.  
XX OS Homo sapiens.  
XX FT Key Location/Qualifiers  
XX FT Misc-difference 56 /note= "Encoded by AAG"  
XX FT Misc-difference 554.572 /note= "Not encoded by the DNA sequence appearing as  
XX FT ABK91191"

US2002119944-A1.



```

PD 29-AUG-2002.
XX
XX
XX 09-NOV-2001; 2001US-00986632.
XX
XX 09-NOV-2000; 2000US-0246751P.
XX
XX (AGUE/) AGUERA M.
XX PA (BELI/) BELIN M.
XX PA (CHAR/) CHARRIER E.
XX PA (HONO/) HONORAT J.
XX PA (RICA/) RICARD D.
XX PA (ROGE/) ROGEMOND V.
XX
XX Aguera M, Belin M, Charrier E, Honorat J, Ricard D, Rogemond V;
XX WPI; 2002-627172/67.
XX
XX Prevention or treatment of myelin disorders, such as multiple sclerosis,
XX PT by administering an agent selected from a Ulip/CRMP protein, a nucleic
XX PT acid coding for the protein, or an antibody directed against protein.
XX
XX
XX Disclosure; Page 26-27; 44pp; English.
XX
XX The invention relates to a new method for prevention or treatment of
XX CC myelin disorders, comprises administering to a patient an effective
XX CC amount of an agent selected from a Ulip (Unc-33-like protein/CRMP
XX CC (collapsin response mediator protein) protein, a nucleic acid coding for
XX CC Ulip/CRMP, an antisense sequence capable of specifically hybridizing with
XX CC the nucleic acid, an antibody directed against Ulip/CRMP, or an aptamer
XX CC capable of binding Ulip/CRMP, and a pharmacologically acceptable carrier.
XX CC Also included are methods of diagnosing a myelin disorder in a subject,
XX CC identifying agents useful for the prevention or treatment of myelin
XX CC disorders, using the Ulip/CRMP proteins/nucleic acids, agents capable of
XX CC modulating the function or expression of the proteins (increasing or
XX CC decreasing), and a method for identifying an endogenous agent as a
XX CC therapeutic target for the prevention or the treatment of myelin
XX CC disorders. The agents are useful for preventing or treating a myelin
XX CC disorder such as multiple sclerosis or HTLV-1 (human T lymphocyte virus
XX CC 1) associated myelopathy and neurodegenerative diseases, Alzheimer's
XX CC disease, paraneoplastic neurodegenerative diseases (PND), autoimmune
XX CC neurodegenerative disorder. Ulip/CRMP proteins are involved in
XX CC the processes of myelination, demyelination and remyelination. Antibodies
XX CC to a Ulip/CRMP protein are useful for diagnosing a myelin disorder. The
XX CC present sequence represents Ulip4/CRMP3 protein
XX
XX Sequence 572 AA;
XX
XX Query Match
XX Best Local Similarity 99.8%; Score 2976; DB 5; Length 572;
XX Matches 571; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
XX
QY 1 MSFGKKSIPRITSDRLIRGGIRVNDQSFYADVHVDEGLIKQIGENLIVPGIKTIDA 60
DB 1 MSFGKKSIPRITSDRLIRGGIRVNDQSFYADVHVDEGLIKQIGENLIVPGIKTIDA 60
QY 61 HGLMVLPGGVVHTRIQMPVLGTPADDFCOGKKAALAGTTMILDHVPDTCVSLAAY 120
DB 61 HGLMVLPGGVVHTRIQMPVLGTPADDFCOGKKAALAGTTMILDHVPDTCVSLAAY 120
QY 121 EQWRERADSAACDYSILHVDITRWHSIKEEALVKEGVNSFLVPMAYKRCQSDSQ 180
DB 121 EQWRERADSAACDYSILHVDITRWHSIKEEALVKEGVNSFLVPMAYKRCQSDSQ 180
QY 181 MYEIPSIIRDLGALAQVHAENGDI VEEQKRLLELIGITGPEGHVLSHPREVEAEVYRAV 240
DB 181 MYEIPSIIRDLGALAQVHAENGDI VEEQKRLLELIGITGPEGHVLSHPREVEAEVYRAV 240
QY 241 TIAQANCPLYTVTKVSKGAADAIQAKRGVVFGPEPTASIGTDSHYVSNKNKAKAA 300
DB 241 TIAQANCPLYTVTKVSKGAADAIQAKRGVVFGPEPTASIGTDSHYVSNKNKAKAA 300
QY 301 FVTSPVPNPDPPTTADHLTCLLSGDLQVTSAGHCTFTTAQKAVGKDNFALIPGTINGIBE 360

```

---

```

DB 301 FVTSPVPNPDPPTTADHLTCLLSGDLQVTSAGHCTFTTAQKAVGKDNFALIPGTINGIBE 360
QY 361 RMSWWEKCVASGKMDNEFVAVTSTNAAKIFNFPRKGRVAVGSDADLVINPKATKII 420
DB 361 RMSWWEKCVASGKMDNEFVAVTSTNAAKIFNFPRKGRVAVGSDADLVINPKATKII 420
QY 421 SAKTHNLNVEYNI FPGVECRGAPAVVISQGRVALEDGKMFVTPGAGRFPVPRKTFPDPVYK 480
DB 421 SAKTHNLNVEYNI FPGVECRGAPAVVISQGRVALEDGKMFVTPGAGRFPVPRKTFPDPVYK 480
QY 481 RIKARNRLAEIHGVPRGLYDGPVHEVMVPAKPGSGAPARASCPCGKISVPPVRLHQSGFS 540
DB 481 RIKARNRLAEIHGVPRGLYDGPVHEVMVPAKPGSGAPARASCPCGKISVPPVRLHQSGFS 540
QY 541 LSGSQADDDHARRTAQKIMAPPGGRSNTISLS 572
DB 541 LSGSQADDDHARRTAQKIMAPPGGRSNTISLS 572
XX
XX RESULT 5
XX AAW68489
XX ID AAW68489 standard; protein; 553 AA.
XX
XX AAW68489;
XX AC
XX DT 08-DEC-1998 (first entry)
XX DE Human partial ULIP-4 protein.
XX KW Human; Unc-33-like phospho-protein; ULIP; rat; neoplasm; tumorigenesis;
XX KW neurodegenerative disorder; diagnosis.
XX OS Homo sapiens.
XX
XX FH Key Location/Qualifiers
XX FT Misc-difference 56 /note= "encoded by TGA"
XX FT
XX FN FR2759701-AL.
XX
XX PD 21-AUG-1998.
XX
XX PF 19-FEB-1997; 97FR-00001961.
XX
XX PR 19-FEB-1997; 97FR-00001961.
XX
XX (INRM ) INSERM INST NAT SANTE & RECH MEDICALE.
XX
XX PI Aguera M, Belin MF, Honnorat J, Kolattukudy P, Quach TT, Byk T;
XX PI Sobel A;
XX
XX WPI; 1998-449610/39.
XX DR N-PSDB; AAV60818.
XX
XX Mouse and human ULIP poly:peptide(s) - useful in detection of para-
XX PT neoplastic neurological syndromes.
XX PS Claim 1; Fig 12; 90pp; French.
XX
XX This sequence represents a partial human Unc-33-like phospho-protein
XX CC (ULIP)-4. The coding sequence was isolated based on similarity to the rat
XX CC ULIP sequence. Proteins of the ULIP family or their corresponding nucleic
XX CC acids can be used in compositions for treating neurodegenerative
XX CC disorders and neoplasms, especially for para-neoplastic neurological
XX CC syndromes and/or for the early diagnosis of tumorigenesis
XX
XX Sequence 553 AA;
XX
XX Query Match 96.6%; Score 2882; DB 2; Length 553;
XX Best Local Similarity 99.8%; Pred. No. 8.9e-268;
XX Matches 552; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
XX
QY 1 MSFGKKSIPRITSDRLIRGGIRVNDQSFYADVHVDEGLIKQIGENLIVPGIKTIDA 60

```

```

Db 1 MSFGKKSIPRITSDRLIRGGRIVNDQSFYADVHVEDGLIKQIGENLIVPGGIKTIDA 60
QY 61 HGLMWLPGGVDVHTRQLQMPVLGTMTPADDFCQGTKAALAGGTTMILDHVPDPTGVSLLAAY 120
Db 61 HGLMWLPGGVDVHTRQLQMPVLGTMTPADDFCQGTKAALAGGTTMILDHVPDPTGVSLLAAY 120
QY 121 EQRERADSAACCDYSLHVDITRWHSIKELEALVKEKGVNSFLVPMAYKDRQCQSDSQ 180
Db 121 EQRERADSAACCDYSLHVDITRWHSIKELEALVKEKGVNSFLVPMAYKDRQCQSDSQ 180
QY 181 MYEIPSIIRDLGALAOVHAENGDIVEEEOKRLLELGTGPEGHVLSHPPEVEAEVYRAV 240
Db 181 MYEIPSIIRDLGALAOVHAENGDIVEEEOKRLLELGTGPEGHVLSHPPEVEAEVYRAV 240
QY 241 TIQAQNCPLVYTKWMSKGAADAIAQAKRGVVVFGPEPTASLGTGSHYWSKNWAKAAA 300
Db 241 TIQAQNCPLVYTKWMSKGAADAIAQAKRGVVVFGPEPTASLGTGSHYWSKNWAKAAA 300
QY 301 FVTSPVPNPDPPTADHLTCLLSSGDLQVTSAGHCTFTTAQKAVGKONFALIPGTTNGIEE 360
Db 301 FVTSPVPNPDPPTADHLTCLLSSGDLQVTSAGHCTFTTAQKAVGKONFALIPGTTNGIEE 360
QY 361 RMSWWEKCVASGKMDENEFVAVTSTNAAKIFNFPYPRKGRVAVGSDADLVINWPKATKII 420
Db 361 RMSWWEKCVASGKMDENEFVAVTSTNAAKIFNFPYPRKGRVAVGSDADLVINWPKATKII 420
QY 421 SAKTHNLNVEYNIFEGVECKGAPAVVISQGRVALEDGKMFVTPGAGRFVPRKTFPPDPVYK 480
Db 421 SAKTHNLNVEYNIFEGVECKGAPAVVISQGRVALEDGKMFVTPGAGRFVPRKTFPPDPVYK 480
QY 481 RIKARNRLAEIHGVPRLYDGPVHEVMVPAKPGSGAPARASCPCGKISVPPVRNLHQS GFS 540
Db 481 RIKARNRLAEIHGVPRLYDGPVHEVMVPAKPGSGAPARASCPCGKISVPPVRNLHQS GFS 540
QY 541 LSGSQADDDHIARR 553
Db 541 LSGSQADDDHIARR 553

RESULT 6
AAW68488 standard; protein; 572 AA.
XX
AC AAW68488;
XX
DT 08-DEC-1998 (first entry)
XX
DE Mouse ULIP-4 protein.
XX
KW Mouse; Unc-33-like phospho-protein; ULIP; rat; neoplasm; tumorigenesis;
KW neurodegenerative disorder; diagnosis.
XX
OS Mus musculus.
XX
PN FR2759701-A1.
XX
PD 21-AUG-1998.
XX
PF 19-FEB-1997; 97FR-00001961.
XX
PR 19-FEB-1997; 97FR-00001961.
XX
PA (INRM ) INSERM INST NAT SANTE & RECH MEDICALE.
XX
PI Aguerre M, Belin MF, Honnorat J, Kolattukudy P, Quach TT, Byk T;
PI Sobel A;
XX
DR WPI; 1998-449610/39.
DR N-PSDB; AAV60817.
XX
PT Mouse and human ULIP poly:peptide(s) - useful in detection of para-
neoplastic neurological syndromes.

```

```

XX
PS Claim 1; Fig 11; 90pp; French.
XX
CC This sequence represents the mouse Unc-33-like phospho-protein (ULIP)-4.
CC The coding sequence was isolated based on similarity to the rat ULIP
CC sequence. Proteins of the ULIP family or their corresponding nucleic
CC acids can be used in compositions for treating neurodegenerative
CC disorders and neoplasms, especially for para-neoplastic neurological
CC syndromes and/or for the early diagnosis of tumorigenesis
XX
SQ Sequence 572 AA;

```

```

Query Match 94.2%; Score 2809; DB 2; Length 572;
Best Local Similarity 92.8%; Pred. No. 1e-260;
Matches 531; Conservative 22; Mismatches 19; Indels 0; Gaps 0;

QY 1 MSFGKKSIPRITSDRLIRGGRIVNDQSFYADVHVEDGLIKQIGENLIVPGGIKTIDA 60
Db 1 MSFGKKSIPRITSDRLIRGGRIVNDQSFYADVHVEDGLIKQIGENLIVPGGIKTIDA 60
QY 61 HGLMWLPGGVDVHTRQLQMPVLGTMTPADDFCQGTKAALAGGTTMILDHVPDPTGVSLLAAY 120
Db 61 HGLMWLPGGVDVHTRQLQMPVLGTMTPADDFCQGTKAALAGGTTMILDHVPDPTGVSLLAAY 120
QY 121 EQRERADSAACCDYSLHVDITRWHSIKELEALVKEKGVNSFLVPMAYKDRQCQSDSQ 180
Db 121 EQRERADSAACCDYSLHVDITRWHSIKELEALVKEKGVNSFLVPMAYKDRQCQSDSQ 180
QY 181 MYEIPSIIRDLGALAOVHAENGDIVEEEOKRLLELGTGPEGHVLSHPPEVEAEVYRAV 240
Db 181 MYEIPSIIRDLGALAOVHAENGDIVEEEOKRLLELGTGPEGHVLSHPPEVEAEVYRAV 240
QY 241 TIQAQNCPLVYTKWMSKGAADAIAQAKRGVVVFGPEPTASLGTGSHYWSKNWAKAAA 300
Db 241 TIQAQNCPLVYTKWMSKGAADAIAQAKRGVVVFGPEPTASLGTGSHYWSKNWAKAAA 300
QY 301 FVTSPVPNPDPPTADHLTCLLSSGDLQVTSAGHCTFTTAQKAVGKONFALIPGTTNGIEE 360
Db 301 FVTSPVPNPDPPTADHLTCLLSSGDLQVTSAGHCTFTTAQKAVGKONFALIPGTTNGIEE 360
QY 361 RMSWWEKCVASGKMDENEFVAVTSTNAAKIFNFPYPRKGRVAVGSDADLVINWPKATKII 420
Db 361 RMSWWEKCVASGKMDENEFVAVTSTNAAKIFNFPYPRKGRVAVGSDADLVINWPKATKII 420
QY 421 SAKTHNLNVEYNIFEGVECKGAPAVVISQGRVALEDGKMFVTPGAGRFVPRKTFPPDPVYK 480
Db 421 SAKTHNLNVEYNIFEGVECKGAPAVVISQGRVALEDGKMFVTPGAGRFVPRKTFPPDPVYK 480
QY 481 RIKARNRLAEIHGVPRLYDGPVHEVMVPAKPGSGAPARASCPCGKISVPPVRNLHQS GFS 540
Db 481 RIKARNRLAEIHGVPRLYDGPVHEVMVPAKPGSGAPARASCPCGKISVPPVRNLHQS GFS 540
QY 541 LSGSQADDDHIARRTAOKIMAPPGGRSNIITSL 572
Db 541 LSGSQADDDHIARRTAOKIMAPPGGRSNIITSL 572

```

```

RESULT 7
AAE38226
ID AAE38226 standard; protein; 512 AA.
XX
AC AAE38226;
XX
DT 20-NOV-2003 (first entry)
XX
DE Human enzyme (ENZYME) protein #18.
XX
KW Human; enzyme; ENZYME; immune disorder; infection; myocardial infarction;
KW gene therapy; anaemia; acquired immune deficiency syndrome; infection;
KW reproductive disorder; cardiovascular; eye; cell proliferation; cancer;
KW AIDS; allergy; asthma; Addison's disease; diabetes; goitre; impotence;
KW infertility; atherosclerosis; metabolic disorder.
XX

```



Db 61 YGQLVVPGGIDVHTRLQMAVKGASADDFYQGTAAALAGTTIMDHVLPAGTSLAAAY 120  
 QY 121 EQRERADSAACDYSLSHVDITWHSISKEELALVKEKGVNSFLVPMAYKDCQCSDSQ 180  
 Db 121 EQRQRADSRACDYLHIIDIPRWHSLELEALVKDGVNSFLVPMAYKDRLOCTDAQ 180  
 QY 181 MYEIFSIRDLGALAQVHAENGDI VEEBQKRLLELIGITGPEGHVLSHPPEVEAEAVYRAV 240  
 Db 181 MYEIFCIIRDLGALAQVHAENGDIIEEBQKRLLDIGITGPEGHVLSRPEVEAEAVYRAI 240  
 QY 241 TIAKQANCPVYTKVMSKGAADAIQAQRKGVVVFGEPIITASIGTDSHYWNSKWAQAA 300  
 Db 241 TIAKQANCPVYTKVMSKGAADVVAQAQRKGVVVFGEPIITASIGADGSHYWSKWAQAA 300  
 QY 301 FVTSPPVNPDPPTADHLTCLLSSGDLQVTSAGHCTFTTAQKAVGKONFALIPGTTNGIBE 360  
 Db 301 FVTSPLSPDPPTPERLSLLSCGDLQVAGSAHCTFTTAQKAVGKONFALIPGTTNGIBE 360  
 QY 361 RMSMWKCVASGKMDNEFVAVTSTNAAKIFNYPKRGVAVGSDADLVIMNPKATKII 420  
 Db 361 RMAIWEKCVSGKMDNEFVAVTSTNAAKIFNYPKRGVAVGSDADLVIMNPKATKVI 420  
 QY 421 SAKTHNLNVEYNIFEGVECEGAPAVVISQGRVALEDGKMFVTPGAGRFVPRKTFPDPVYK 480  
 Db 421 SAKTHNLNVEYNIFEGTECHGAPAVVISQGVVLEDNLFVTEGSGRFVPRKTFPDPVYK 480  
 QY 481 RIKARNRLAEIHGVPRGLYDGPVHEVWVPAK--PGSGAPARASCPGKISVPPVRNLHOSG 538  
 Db 481 RIKARNRLAEVHGVPRGLYDGPVHDVVLSTKAVPTTAAASRIACAGKVPAPVRNLHOSG 540  
 QY 539 FSLSGSQADHDHARRTAQKIMAPPGGRSNTISLS 572  
 Db 541 FSLSGSQADHDHARRTAQKIMAPPGGRSNTISLS 574

RESULT 9  
 ID ADK70696  
 XX ADK70696 standard; protein; 650 AA.  
 AC ADK70696;  
 XX  
 DT 06-MAY-2004 (first entry)  
 XX Chicken CRMP3A protein isoform SeqID.  
 DE signal control molecule; collapsin response mediator protein; CRMP;  
 KW nerve growth cone; axial future edge; neural-network formation;  
 KW cell polarity; neurogenesis; chicken.  
 XX  
 OS Gallus gallus.  
 XX  
 PN JP2004000094-A.  
 XX  
 PD 08-JAN-2004.  
 XX  
 PF 31-MAY-2002; 2002JP-00160853.  
 XX  
 PR 31-MAY-2002; 2002JP-00160853.  
 XX  
 PA (KAGA-) KAGAKU GIJUTSU SHINKO JIGYODAN.  
 XX  
 DR WPI; 2004-085208/09.  
 DR N-PSDB; ADK70695.  
 XX  
 PT Novel N-terminal variant polypeptide of signal control molecule collapsin  
 PT response mediator protein of nerve growth cone, useful for neural-network  
 PT formation.  
 XX  
 PS Claim 14; SEQ ID NO 6; 94pp; Japanese.  
 XX  
 CC This invention relates to a novel variant of an intracellular signal  
 CC control molecule identified as the collapsin response mediator protein

CC (CRMP) that works within the steering mechanism of the nerve growth cone  
 CC in an axial future edge part. Specifically, it refers to an N-terminal  
 CC variant that exhibits a deletion, substitution or addition of an amino  
 CC acid in CRMP, as well as the appropriate monoclonal antibody. The present  
 CC invention describes the CRMP variant as useful for neural-network  
 CC formation and for the establishment of cell polarity during growth.  
 CC Furthermore, it can be efficiently used during the neurogenesis process.  
 CC This polypeptide sequence is a chicken CRMP isoform of the invention.  
 XX  
 SQ Sequence 650 AA;

Query Match 83.4%; Score 2488; DB 8; Length 650;  
 Best Local Similarity 81.8%; Pred. No. 1e-229;  
 Matches 459; Conservative 56; Mismatches 44; Indels 2; Gaps 1;

QY 14 SDRLLIRGRIVNDQSFYADVHVRDGLIKOIGENLIVPGGIKTIIDAHGLMWLPGGVVHV 73  
 Db 90 SQOLLIRGRIVNDQSFYADVHVRDGLIKOIGENLIVPGGVVHV 149  
 QY 74 TRLQMPVLGMPADDFCQGTAAALAGTTIMDHVLPAGTSLAAAYEQWRQADSRACC 133  
 Db 150 TRLQMAVMSGASADDFYQGTAAALAGTTIMDHVLPAGTSLAAAYEQWRQADSRACC 209  
 QY 134 DYSLHVDITRWHSISKEELALVKEKGVNSFLVPMAYKDCQCSDSQMTFPIIRDLGA 193  
 Db 210 DYALHIDIIRWHSLELEALVKDGVNSFLVPMAYKDRLOCTDAQMTFPIIRDLGA 269  
 QY 194 LAQVHAENGDI VEEBQKRLLELIGITGPEGHVLSHPPEVEAEAVYRAVTTAKQANCPVYT 253  
 Db 270 IAQVHAENGDIIEEBQKRLLDIGITGPEGHVLSRPEVEAEAVYRAITITAKQANCPVYT 329  
 QY 254 KYMSKGAADAIQAQRKGVVVFGEPIITASIGTDSHYWNSKWAQAAAFVTSPPVNPDPPT 313  
 Db 330 KYMSKGAADVVAQAQRKGVVVFGEPIITASIGADGSHYWSKWAQAAAFVTSPLSPDPPT 389  
 QY 314 ADHLTCLLSSGDLQVTSAGHCTFTTAQKAVGKONFALIPGTTNGIBEERMSMWKCVASG 373  
 Db 390 PERLSLLSCGDLQVAGSAHCTFTTAQKAVGKONFALIPGTTNGIBEERMAIWEKCVSG 449  
 QY 374 KMDNEFVAVTSTNAAKIFNYPKRGVAVGSDADLVIMNPKATKII SAKTHNLNVEYNI 433  
 Db 450 KMDNEFVAVTSTNAAKIFNYPKRGVAVGSDADLVIMNPKATKIVISAKTHNLNVEYNI 509  
 QY 434 FEGVECEGAPAVVISQGRVALEDGKMFVTPGAGRFVPRKTFPDPVYKRIKARNRLAEIHG 493  
 Db 510 FEGTECHGAPAVVISQGVVLEDNLFVTEGSGRFVPRKTFPDPVYKRIKARNRLAEVHG 569  
 QY 494 VPRGLYDGPVHEVWVPAK--PGSGAPARASCPGKISVPPVRNLHOSGFSLSGQADHDHIA 551  
 Db 570 VPRGLYDGPVHDVVLSTKAVPTTAAASRIACAGKVPAPVRNLHOSGFSLSGQADHDHVA 629  
 QY 552 RETAKKIMAPPGGRSNTISLS 572  
 Db 630 RRTAKKIMAPPGGRSNTISLS 650

RESULT 10  
 ID ADK70707 standard; protein; 572 AA.  
 XX ADK70707;  
 AC ADK70707;  
 XX  
 DT 06-MAY-2004 (first entry)  
 XX Collapsin response mediator protein (CRMP) 2B SeqID.  
 DE signal control molecule; collapsin response mediator protein; CRMP;  
 KW nerve growth cone; axial future edge; neural-network formation;  
 KW cell polarity; neurogenesis.  
 XX  
 OS Unidentified.  
 XX  
 PN JP2004000094-A.



Matches 434; Conservative 70; Mismatches 67; Indels 0; Gaps 0;	
Qy 1	MSPGQKKSIPRITSDRLIRGGRIVNDQSFYADVHVEDGLIKQIGENLIVPGGIKTIDA 60
Db 1	MSYQKKNIIPRITSDRLIRGGKI VNDQSFYADYIMEDGLIKQIGENLIVPGVKITIEA 60
Qy 61	HGLMVLPGGVVHTRIQMPVLGHTPADDPCQGTAKAALAGGTTMILDHVPDGTGVSLLAAY 120
Db 61	HSRMVTPGGIDVHTRFQMPDQGTSAADFFQGTAKAALAGGTTMIDHVPDPGTGSLAAP 120
Qy 121	EQWREADSAACDYSLSHVDITRWHSIEKELEALVKEKGVNSFLVFMVKDRCCQSDSQ 180
Db 121	DQWREADSKCCDYSLSHVDITEWHKIGIQEEMALVKDHGVNSFLVFMVKDRFQLTDSQ 180
Qy 181	MYBIFSIIRDGLAQAQVHAENGDI VEEQKRLLELGTGPEGHVLSHPPEVEAEAVYRAV 240
Db 181	IYEVLSVIRDIGAIAQVHAENGDI IAEQQRILDLGITGPEGHVLSRPEVEAEAVNRSI 240
Qy 241	TIKQANCPLYVTVMKSKGAADIAQAQRGVVVFGEPI TASIGTDSHYWKNWAKAAA 300
Db 241	TIANQTNCPLYVTVMKSKSAARVIAQAQRKGT VYGEPI TASIGTDSHYWKNWAKAAA 300
Qy 301	FVTSPPVNPDPPTADHLTCLLSGDLQVTSAGHCTFTTAOKAVGKDNFALIPGTNGIEE 360
Db 301	FVTSPLSPDPPTDFLNSLLSCGDLQVTSAGHCTFTTAOKAVGKDNFTLIPGTNGTEE 360
Qy 361	RMSMWKCVASGKMDENEFVAVTSTNAAKIFNFPYPRKGRVAVGSDADLVINPKATKII 420
Db 361	RMSVIMDKAVVTGKMDENQFVAVTSTNAAKVFNLYPRKGRISVGSDADLVIMPDVKTI 420
Qy 421	SAKTHNLNVEYNIPEGVECRGAPAVVISQGRVALEDGKMFVTPGAGRFPVPRKTFPDPVYK 480
Db 421	SAKTHNSALEYNIPEGMECRGSLVVISQKIVLEDTLHVTEGSGRYIPRKEFPDPVYK 480
Qy 481	RIKARNRLAEIHGVPRGLYDGPVHEVMVPAKPGSGAPARASCPGKISVPPVRNLHSGGFS 540
Db 481	RIKARSRLAELRGVPRGLYDGPVCEVSVTPKTVTTPASSAKTSPAKQOAPPVRNLHSGGFS 540
Qy 541	LSGSQADDDHIARTAKIMAPPGGRSNTSL 571
Db 541	LSGAQIDDDNI PRRTQRI VAPPGGRANITSL 571
RESULT 12	
ID	ADJ84279
XX	standard; protein; 572 AA.
AC	ADJ84279;
XX	
DT	06-MAY-2004 (first entry)
XX	
DE	Malayan black rat CRMP-2 (collapsein response mediator protein 2) protein.
XX	
KW	antimicrobial; antiinflammatory; cytostatic; infection; inflammation;
KW	tumour formation; antisense therapy; Malayan black rat; CRMP-2;
XX	
OS	Rattus rattus.
XX	
PN	WO2004003134-A2.
XX	
PD	08-JAN-2004.
XX	
PF	12-JUN-2003; 2003WO-US018481.
XX	
PR	26-JUN-2002; 2002US-0392020P.
XX	
PA	(ISIS-) ISIS PHARM INC.
XX	
PI	Monia Bp; Freier SM, Manoharan M, Gaarde WA;
XX	
DR	WPI; 2004-083026/08.
DR	N-PSDB; ADJ84229.

XX

PT

PT

PT

PT

XX

PS

XX

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

XX

SQ

```

XX 05-NOV-2002 (first entry)
DT Human Ulp2/CRMP2 protein.
XX
XX Human; Ulp2; CRMP2; collapsin response mediator protein;
XX Unc-33-like protein; neurodegenerative disease; Alzheimer's disease;
XX paraneoplastic neurodegenerative disease; PND; myelination;
XX demyelination; remyelination; myelin disorder; multiple sclerosis;
XX autoimmune neurodegenerative disorder; HTLV-1 associated myelopathy;
XX human T lymphocyte virus 1.
XX
OS Homo sapiens.
XX
XX US2002119944-A1.
XX
XX 29-AUG-2002.
XX
XX 09-NOV-2001; 2001US-00986632.
XX
XX 09-NOV-2000; 2000US-0246751P.
XX
XX (AGUE/) AGUERA M.
XX (BELI/) BELIN M.
XX (CHAR/) CHARRIER E.
XX (HONO/) HONORAT J.
XX (RICA/) RICARD D.
XX (ROGE/) ROGEMOND V.
XX
XX Aguera M, Belin M, Charrier E, Honorat J, Ricard D, Rogemond V;
XX WPI; 2002-627172/67.
XX N-PSDB; ABK91188.
XX
XX Prevention or treatment of myelin disorders, such as multiple sclerosis,
XX by administering an agent selected from a Ulp1/CRMP protein, a nucleic
XX acid coding for the protein, or an antibody directed against protein.
XX
XX Claim 12; Page 18-20; 44pp; English.
XX
XX The invention relates to a new method for prevention or treatment of
XX myelin disorders, comprises administering to a patient an effective
XX amount of an agent selected from a Ulp1 (Unc-33-like protein)/CRMP
XX (collapsin response mediator protein) protein, a nucleic acid coding for
XX Ulp1/CRMP, an antisense sequence capable of specifically hybridising with
XX the nucleic acid, an antibody directed against Ulp1/CRMP, or an aptamer
XX capable of binding Ulp1/CRMP, and a pharmacologically acceptable carrier.
XX Also included are methods of diagnosing a myelin disorder in a subject,
XX identifying agents useful for the prevention or treatment of myelin
XX disorders, using the Ulp1/CRMP proteins/nucleic acids, agents capable of
XX modulating the function or expression of the proteins (increasing or
XX decreasing), and a method for identifying an endogenous agent as a
XX therapeutic target for the prevention or the treatment of myelin
XX disorders. The agents are useful for preventing or treating a myelin
XX disorder such as multiple sclerosis or HTLV-1 (human T lymphocyte virus
XX 1) associated myelopathy and neurodegenerative diseases, Alzheimer's
XX disease, paraneoplastic neurodegenerative diseases (PND), autoimmune
XX neurodegenerative disorder. Ulp1/CRMP proteins are involved in
XX the processes of myelination, demyelination and remyelination. Antibodies
XX to a Ulp1/CRMP protein are useful for diagnosing a myelin disorder. The
XX present sequence represents Ulp2/CRMP2 protein
XX
XX Sequence 572 AA;
XX
XX Query Match 78.6%; Score 2345; DB 5; Length 572;
XX Best Local Similarity 75.7%; Pred. No. 4.8e-216;
XX Matches 432; Conservative 72; Mismatches 67; Indels 0; Gaps 0;
XX
XX 1 MSFGKSTPRITSDRLIRGGHIVNDQSFYADVHVEDGLIKQIGENLVPGIKTIDA 60
XX :||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
XX 1 MSYQGGKNTPRITSDRLIRGGHIVNDQSFYADYIMEDGLIKQIGENLVPGVGTIEA 60
XX
XX 61 HGLMWLPVGGVDVHTRLQMPVLGHTPADDFCQGTAKAALAGTTMILDHVPDTCVSLAAY 120

```

```

Db 61 HSRWIPGGIDVHTRFQMPDQMTSADDFQGTAKAALAGTTMILDHVPDTCVSLAAY 120
QY 121 EQWRERADSAACDYSYLVHVDITRMHESIKKELEALVKEGVNSFLVFMAYKDRCCQSDSQ 180
Db 121 DQWRERADSAACDYSYLVHVDITRMHESIKKELEALVKEGVNSFLVFMAYKDRCCQSDSQ 180
QY 181 MYEIPSIIRIDICALAQVHAENGDIIVEEQKRLLELIGITGPEGHVLSHPPEVEAEAYRAV 240
Db 181 IYEVLSVIRIDICALAQVHAENGDIIVEEQKRLLELIGITGPEGHVLSHPPEVEAEAVNRAI 240
QY 241 TIAKQANCPYVTKVMSKGAADAIAQAKRRGVVVFGEPIITASLTGTDGSHYMSKNWAKAAA 300
Db 241 TIANQTCCLPYITKWSKSSAFAVIAQAKRGKGVVVFGEPIITASLTGTDGSHYMSKNWAKAAA 300
QY 301 FVTSPVPNPDPPTADHLTCLSSGDLQVTGSAHCTFTTAQKAVGKDNFALIPETNGIEE 360
Db 301 FVTSPPLSPDPPTADHLTCLSSGDLQVTGSAHCTFTTAQKAVGKDNFALIPETNGIEE 360
QY 361 RMSVWEKCVASGKMDENEFVAVTSTNAKIFNFPYKRGVAVGSDADLVINPKATKII 420
Db 361 RMSVWDXKAVVTGKMDENEFVAVTSTNAKIFNFPYKRGVAVGSDADLVINPKATKII 420
QY 421 SAKTNLNVYNIIFEGVECRGAPAVVISQCRVALEDGKMFVTPGAGRFPVPRKTFPDPVVK 480
Db 421 SAKTNSSLEYNIIFEGVECRGAPAVVISQCRVALEDGKMFVTPGAGRFPVPRKTFPDPVVK 480
QY 481 RIKARNRLAEIHGVPRLYDGPVHEVMVPAKPGSGAPARASCPGKISVPPVNLHQSGFS 540
Db 481 RIKARSRLAEIHGVPRLYDGPVHEVMVPAKPGSGAPARASCPGKISVPPVNLHQSGFS 540
QY 541 LSGSQADDDHARRTAQKIMAPGGRSNTSL 571
Db 541 LSGAQDDDDHARRTAQKIMAPGGRSNTSL 571
XX
XX RESULT 14
XX ABP97908
XX ID ABP97908 standard; protein; 572 AA.
XX
XX AC ABP97908;
XX
XX DT 17-JUN-2003 (first entry)
XX
XX DE Amino acid sequence of human CRMP-2.
XX
XX KW CRMP; collapsin response mediator protein; T lymphocyte; T cell;
XX immune system disorder; prion protein; T cell leukemia; T cell lymphoma;
XX viral infection; prion disease; demyelinating neuroinflammatory disease;
XX multiple sclerosis; prion disease; CRMP-2.
XX
XX OS Homo sapiens.
XX
XX PN WO2003022298-A2.
XX
XX PD 20-MAR-2003.
XX
XX PF 09-SEP-2002; 2002WO-FR003056.
XX
XX PR 07-SEP-2001; 2001FR-00011627.
XX PR 16-OCT-2001; 2001FR-00013342.
XX
XX PA (INRM ) INSERM INST NAT SANTE & RECH MEDICALE.
XX
XX PI Giraudon P, Belin M, Malcus C, Colas P, Antoine J, Honnorat J;
XX WPI; 2003-301008/29.
XX
XX PT Use of collapsin response mediator protein for treating T lymphocyte
XX dysfunction, e.g. viral infection or leukemia, also for drug screening,
XX diagnosis and prognosis.
XX
XX PS Disclosure; Fig 8; 58pp; French.

```





Db	421	SAKTHNSSLEYNIFEGMECRGSPVVISQKIVLEDTLHVTEGSGRYIPRKPPDFVYK	480
Qy	481	RIKARNRLAEIHCVPRGLYDGPVHEVMVPAPKPCSGAPASCEGKISVPVRLHQS	540
Db	481	RIKARSRLAELRGVPRGLYDGPVCEVSVTPKVTTPASSAKTSPAKQOQAPPVRLHQS	540
Qy	541	LSGSQADDDHARRTAQKIMAPPGRSNIITSL	571
Db	541	LSGAQIDDNIPTTQRIVAPPGGRANITSL	571

Search completed: September 24, 2005, 17:11:29  
Job time : 1138 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 24, 2005, 16:34:26 ; Search time 59 Seconds  
(without alignments)  
723.716 Million cell updates/sec

Title: US-09-367-496C-8  
Perfect score: 2982  
Sequence: 1 MSFQKKSIPRITSDRLIR.....RTAQKIMAPPGGRSNTISLS 572

Scoring table:  
BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents AA:\*  
1: /cgn2.6/prodata/1/iaa/5A COMB.pap.\*  
2: /cgn2.6/prodata/1/iaa/5B COMB.pap.\*  
3: /cgn2.6/prodata/1/iaa/6A COMB.pap.\*  
4: /cgn2.6/prodata/1/iaa/6B COMB.pap.\*  
5: /cgn2.6/prodata/1/iaa/6C COMB.pap.\*  
6: /cgn2.6/prodata/1/iaa/backfiles1.pap.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2345	78.6	572	4	US-09-949-016-6070
2	2345	78.6	578	4	US-09-949-016-11503
3	2152	73.5	570	4	US-09-949-016-6071
4	2152	73.5	593	4	US-09-949-016-10192
5	2157	72.3	572	4	US-09-702-705-1815
6	2157	72.3	572	4	US-09-736-457-1815
7	2157	72.3	572	4	US-09-671-325-1815
8	2157	72.3	622	4	US-09-949-016-10606
9	1074.5	36.0	470	4	US-09-949-016-11197
10	1074.5	36.0	507	4	US-09-252-991A-23541
11	1068	35.8	484	4	US-10-114-810-2
12	1041	34.9	495	4	US-09-734-237B-62
13	1041	34.9	496	4	US-09-734-237B-64
14	827	27.7	460	1	US-08-289-709-1
15	827	27.7	460	1	US-08-602-656-1
16	809	27.1	461	4	US-09-836-992-1
17	699.5	23.5	457	4	US-09-134-000C-6461
18	693.5	23.3	457	4	US-08-415-658-21
19	538.5	18.1	297	4	US-09-270-767-32995
20	538.5	18.1	297	4	US-09-270-767-48212
21	427	14.3	459	4	US-09-950-772-4
22	404.5	13.6	458	4	US-09-407-062-9
23	398.5	13.4	458	4	US-09-497-585A-8
24	398.5	13.4	458	4	US-09-497-585A-14
25	398.5	13.4	458	4	US-09-916-501A-14
26	398.5	13.4	458	4	US-09-916-501A-20
27	397.5	13.3	458	4	US-09-497-585A-4

28	397.5	13.3	458	4	US-09-916-501A-10	Sequence 10, Appl
29	396.5	13.3	458	4	US-09-497-585A-2	Sequence 2, Appl
30	396.5	13.3	458	4	US-09-497-585A-10	Sequence 10, Appl
31	396.5	13.3	458	4	US-09-497-585A-12	Sequence 12, Appl
32	396.5	13.3	458	4	US-09-916-501A-8	Sequence 8, Appl
33	396.5	13.3	458	4	US-09-916-501A-16	Sequence 16, Appl
34	396.5	13.3	458	4	US-09-916-501A-18	Sequence 18, Appl
35	392.5	13.2	458	4	US-09-497-585A-6	Sequence 6, Appl
36	392.5	13.2	458	4	US-09-916-501A-12	Sequence 12, Appl
37	311.5	10.4	454	4	US-09-134-000C-5977	Sequence 1977, Ap
38	262	8.8	411	4	US-09-270-767-33495	Sequence 33495, A
39	242.5	8.1	422	4	US-09-583-110-3849	Sequence 3849, Ap
40	242.5	8.1	426	4	US-09-107-433-3349	Sequence 3349, Ap
41	242	8.1	448	4	US-09-134-000C-6087	Sequence 6087, Ap
42	232	7.8	430	4	US-09-107-532A-5107	Sequence 5107, Ap
43	227	7.6	425	4	US-09-710-279-2604	Sequence 2604, Ap
44	227	7.6	427	3	US-09-134-001C-4803	Sequence 4803, Ap
45	212.5	7.1	734	4	US-09-902-540-11399	Sequence 11399, A

ALIGNMENTS

RESULT 1

US-09-949-016-6070  
; Sequence 6070, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 6070  
; LENGTH: 572  
; TYPE: PRT  
; ORGANISM: Human  
US-09-949-016-6070

Query Match 78.6%; Score 2345; DB 4; Length 572;  
Best Local Similarity 75.7%; Pred. No. 1.8e-237;  
Matches 432; Conservative 72; Mismatches 67; Indels 0; Gaps 0;

QY	1	MSFQKKSIPRITSDRLIRGGRIYNDQSFYADVHVEDGLIKQIGENLIVPGIKTIDA 60	
DB	1	MSYQKKNIPRITSDRLIRGGRIYNDQSFYADVHVEDGLIKQIGENLIVPGIKTIEA 60	
QY	61	HGLMVLPGGVVDVHTRLQMPVLGTPADDFCQGTQKALAGTTLMDHVPDPTGVSLAAY 120	
DB	61	HSEWIPGSDVHTFQMPDQGTADDFCQGTQKALAGTTLMDHVPDPTGVSLAAY 120	
QY	121	EQWRERASAAACDYSYLVHVDITRWHEISKEELALVKEGVNSFLVFMAYKORCQSDSQ 180	
DB	121	DQWRERASAAACDYSYLVHVDITRWHEISKEELALVKEGVNSFLVFMAYKORCQSDSQ 180	
QY	181	MYEISFIRDLALQVHAENGDIVEEQRKLELIGITGPEGHVLSHPREVEAEAYRAV 240	
DB	181	IYEVLISVIRDLALQVHAENGDIIVEEQRKLELIGITGPEGHVLSHPREVEAEAYRAV 240	
QY	241	TIQAQNCPLVYTKWMSKGAADIAQAQRGVVVFGEPIASLGTDGSHYWSKNWAKAAA 300	
DB	241	TIQAQNCPLVYTKWMSKGAADIAQAQRGVVVFGEPIASLGTDGSHYWSKNWAKAAA 300	
QY	301	FVTSPPVNPDPPTTAADHLTCLSSGDLQVTGSAHCTFTTAQKAVGKDNFALIPETGNGIBE 360	

```

Db 301 FVTSPLSPDPTTDFLNSLLSCGDLQVTSAGHCTFNTAQAVGKDNFTLIPBGTNGTSE 360
Qy 361 RSMVMWEKCVASGKMDENEFVAVTSTNAAKIFNFYPRKGRVAVGSDADLVINPKATKII 420
Db 361 RSMVIWDKAVVTGKMDENQFVAVTSTNAAKVFNLYPRKGRIAVGSDADLVINWDPDSVKTI 420
Qy 421 SAKTHNLNVEYNIPEGVCRGAPAVVISQGRVVALEDEGKMFVTPGAGRFPVPRKTFPDPFVYK 480
Db 421 SAKTHNSLSLEYNIFEGMECRGSPLVVISQKIVLEDTGLHVTGSGRYIPRPFDPFVYK 480
Qy 481 RIKARNRLAEIHGVPRGLYDGPVHEVMVPAKPGSGAPARASCPGKISVPPVRLNHQSGFS 540
Db 481 RIKARSRLAELRGVPRGLYDGPVCEVSVTPKTVTTPASSAKTSPAKQQAAPPVRLNHQSGFS 540
Qy 541 LSGSQADDDHARTTAQKIMAPPGRSNITSL 571
Db 541 LSGAQIDDDNIPRRTTORIVAPPGGRANITSL 571

```

## RESULT 2

```

US-09-949-016-11503
; Sequence 11503, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11503
; LENGTH: 578
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-11503

```

```

Query Match 78.6%; Score 2345; DB 4; Length 578;
Best Local Similarity 75.7%; Pred. No. 1.9e-237;
Matches 432; Conservative 72; Mismatches 67; Indels 0; Gaps 0;

Qy 1 MSFQKKSIPRITSDRLIRGRIVNDQSFYADVHVHEDGLIKQIGENLIVPGGIKTIDA 60
Db 7 MSYQKKNIPRITSDRLIKGGKIVNDQSFYADYIMEDGLIKQIGENLIVPGGVKTIEA 66
Qy 61 HGLMVLPGVDVHTRLQMPVLGWTADDFCQGTKAALAGTTMILDHVPDPTGVSLAAY 120
Db 67 HERNVLPGIDVHTRFQMPDQGMTSADDFCQGTKAALAGTTMILDHVPDPTGVSLAAY 126
Qy 121 EOWREPADSAACDYSLSHVDITRWHSIKEELEALVKEGVNSFLVFMAYKDRCCQSDSQ 180
Db 127 DQWREWADSKSCDYSLSHVDISEMHKIQISEMEALVKDHGVNSFLVFMAYKDRFQLTDCQ 186
Qy 181 MYEIPSIIRDLGALQVHAENGDIIVEEQRKLELIGITGPEGHVLSHPREVEAEAYRAV 240
Db 187 IYEVLSVIRDIAGIAQVHAENGDIIBEQOIRLDLIGITGPEGHVLSRPEVEAEAVNRAI 246
Qy 241 TIAKQANCPLYTKVMSKGAADIAQAKRGVVVFGEPI TASLGTGSHYMSKNWAKAAA 300
Db 247 TIANQNCPLLYTKVMSKSAEVAIAQARKKGTVVYGEPI TASLGTGSHYMSKNWAKAAA 306
Qy 301 FVTSPPVNPDPPTADHLTCLLSGDLQVTSAGHCTFNTAQAVGKDNFALIPBGTNGIER 360
Db 307 FVTSPLSPDPTTDFLNSLLSCGDLQVTSAGHCTFNTAQAVGKDNFTLIPBGTNGTSE 366

```

```

Qy 361 RSMVMWEKCVASGKMDENEFVAVTSTNAAKIFNFYPRKGRVAVGSDADLVINPKATKII 420
Db 367 RSMVIWDKAVVTGKMDENQFVAVTSTNAAKVFNLYPRKGRIAVGSDADLVINWDPDSVKTI 426
Qy 421 SAKTHNLNVEYNIPEGVCRGAPAVVISQGRVVALEDEGKMFVTPGAGRFPVPRKTFPDPFVYK 480
Db 427 SAKTHNSLSLEYNIFEGMECRGSPLVVISQKIVLEDTGLHVTGSGRYIPRPFDPFVYK 486
Qy 481 RIKARNRLAEIHGVPRGLYDGPVHEVMVPAKPGSGAPARASCPGKISVPPVRLNHQSGFS 540
Db 487 RIKARSRLAELRGVPRGLYDGPVCEVSVTPKTVTTPASSAKTSPAKQQAAPPVRLNHQSGFS 546
Qy 541 LSGSQADDDHARTTAQKIMAPPGRSNITSL 571
Db 547 LSGAQIDDDNIPRRTTORIVAPPGGRANITSL 577

RESULT 3
US-09-949-016-6071
; Sequence 6071, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6071
; LENGTH: 570
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-6071

```

```

Query Match 73.5%; Score 2192; DB 4; Length 570;
Best Local Similarity 70.1%; Pred. No. 2.4e-221;
Matches 401; Conservative 91; Mismatches 78; Indels 2; Gaps 2;

Qy 1 MSFQKKSIPRITSDRLIRGRIVNDQSFYADVHVHEDGLIKQIGENLIVPGGIKTIDA 60
Db 1 MSYQKKNIPRITSDRLIKGGKIVNDQSFYADYIMEDGLIKQIGENLIVPGGVKTIEA 60
Qy 61 HGLMVLPGVDVHTRLQMPVLGWTADDFCQGTKAALAGTTMILDHVPDPTGVSLAAY 120
Db 61 NGKMWIPGIDVHTRFQMPYKGMTTVDFFCQGTKAALAGTTMILDHVPDPTGVSLAAY 120
Qy 121 EOWREPADSAACDYSLSHVDITRWHSIKEELEALVKEGVNSFLVFMAYKDRCCQSDSQ 180
Db 121 EKWREWADSKSCDYSLSHVDITHWDSVQKQVONLIKQKGVNSFMVYMAKOLYQVSNTE 180
Qy 181 MYEIPSIIRDLGALQVHAENGDIIVEEQRKLELIGITGPEGHVLSHPREVEAEAYRAV 240
Db 181 LYEIFTCLELGAIAQVHAENGDIIAEQETRMLENGITGPEGHVLSRPELEAEAVFRAI 240
Qy 241 TIAKQANCPLYTKVMSKGAADIAQAKRGVVVFGEPI TASLGTGSHYMSKNWAKAAA 300
Db 241 TIASQNCPLLYTKVMSKSAADLISQARKKGNVFGPEPI TASLIGIDGTHYMSKNWAKAAA 300
Qy 301 FVTSPPVNPDPPTADHLTCLLSGDLQVTSAGHCTFNTAQAVGKDNFALIPBGTNGIER 360
Db 301 FVTSPLSPDPTTDPYINSLASGDLQLSGSAHCTFSTAQAKIGNFTAIPEGTNGVEE 360
Qy 361 RSMVMWEKCVASGKMDENEFVAVTSTNAAKIFNFYPRKGRVAVGSDADLVINPKATKII 420
Db 361 RSMVIWDKAVATGKMDENQFVAVTSTNAAKIFNFYPRKGRISVGSDDLVINWDPDAVKIV 420

```





Qy	481	RIKARNLLAEIHGVPRGLYDGVHVVWVPAKPGSGAPARASC	PGKISLVPVPTNLHQSGFS	540
Db	481	EVKLRNKFVGGVSRGMVDGVIEVATPKYATPAPSAKSPKHQPP	PIRNLHQSNFS	540
Qy	541	LSGSAQDHHIARRTAQKIMAPPGGRSNTISL		571
Db	541	LSGAQIDNNPRRTGHRIVAPPPGGRSNTISL		571

```

RESULT 8
US-09-949-016-10606
; Sequence 10606, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10606
; LENGTH: 622
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-10606

```

**Qy** 541 LSGSQADDDHARRTAQKIMAPPGRSNTSL 571  
||||| : ||| : |||||  
**Db** 591 LSGAQIDNNPRRTGHRIVAPPGRSNTSL 621

RESULT 9  
US-09-949-016-11197  
; Sequence 11197, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 11197  
; LENGTH: 470  
; TYPE: PRT  
; ORGANISM: Human  
US-09-949-016-11197

RESULT 10  
US-09-252-991A-23541  
; Sequence 23541, Application US/09252991A  
; Patent No. 6551795  
; GENERAL INFORMATION:  
; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND A  
; TITLE OF INVENTION: ABRUGNOSA FOR DIA  
; FILE REFERENCE: 107196.136



		CURRENT APPLICATION NUMBER: US/09/252,991A			
		PRIOR FILING DATE: 1999-02-18			
		PRIOR APPLICATION NUMBER: US 60/074,788			
		PRIOR FILING DATE: 1998-02-18			
		PRIOR APPLICATION NUMBER: US 60/094,190			
		PRIOR FILING DATE: 1998-07-27			
		NUMBER OF SEQ ID NOS: 33142			
		SEQ ID NO 23541			
		LENGTH: 507			
		TYPE: PRT			
		ORGANISM: Pseudomonas aeruginosa			
		US-09-252-991A-23541			
		Query Match 36.0%; Score 1074.5; DB 4; Length 507;			
		Best Local Similarity 45.8%; Pred. No. 9.2e-104;			
		Matches 216; Conservative 81; Mismatches 166; Indels 9; Gaps 4;			
QY	17	LLIRGRIVNDQSFYADVHVHEDGLIKQIGENLIVPGIKTIDAHGLMVLPGGVVHTRL	76		
DB	31	LLIRGATVTHESYRADVLCAQLGALQIGENLIVPGIKTIDAHGLMVLPGGVVHTRL	90		
QY	77	QMPVLGMPADDFCQGTKAALAGTTTMDLHDVPPDTGVSLLAAYEOWRRERADSAACDYS	136		
DB	91	QLPFGMTVASEDPFSGTAAGLAGTTTSIIDFVIEPNRQSLLEAFHTWRGWAQKSA-ADYG	149		
QY	137	LHVDITRWHESIKERLEALVKEKGVNSFLVPMAYKORCQSDSQMYEIPSIIRDLGALAQ	196		
DB	150	FHVAITWMSDEVARENGELVAQGVNSFKHFMAYKNAIMADTLVASPERCLQLGAVPT	209		
QY	197	VHAENGDIIVEEQKRILLELIGITGPEGHVLSHPREVEAEAVYRAVTIAKQANCLPLYVTKVM	256		
DB	210	VHAENGELVHLQQLLAQLGTLGPEAHLSPRPQVGEAASRAIRIAETLGTPLYLHVIS	269		
QY	257	SKGAADAIQAQKRGVVVFGPEITASIGTDSHYWKNWAKAAAFVTPSPVNPDPPTADH	316		
DB	270	SREALDEIAYARAKQPVYGEVLGHLGLDSDVYRHPDMATAAGYVMSPPFR----	325		
QY	317	LTCL---LQSGDLQVTSAGHCTTTAAQVGNFALIPGTTNGIERMSMWKEKCVASG	373		
DB	326	QEAALWGLQSGNLTATDHCFCASQKAWGRDDFSKIPNGTAGIEDRMALLMDAGVNSG	385		
QY	374	KMDENFVAVTSTNAAKIFNFPYPRKGRVAVGSDADLVINWPKATKIISAKTHNLNVEYNI	433		
DB	386	RLSMHEFVALTSTNAKIFNLPYPRKGRVAVGADLVMDPQGSRTLSAATHQVDFNI	445		
QY	434	FEVGEGRGAPAVVISQGRVALEDGKMFVTPGAGRFVPRKTPDFVYKRIKAR	485		
DB	446	FEGRTVRGIPSHSTISQKLLMAAGDLRAEPGAGRYVERPAYPS-VYEVLGRR	496		
RESULT 11					
		US-10-114-810-2			
		Sequence 2, Application US/10114810			
		Patent No. 6800465			
		GENERAL INFORMATION:			
		APPLICANT: Bristol-Myers Squibb Company			
		TITLE OF INVENTION: D-Hydantoinase From Ochrobactrum anthropi			
		FILE REFERENCE: ON0158-NP			
		CURRENT APPLICATION NUMBER: US/10/114,810			
		PRIOR FILING DATE: 2002-08-16			
		PRIOR APPLICATION NUMBER: US 60/281,150			
		PRIOR FILING DATE: 2001-04-03			
		NUMBER OF SEQ ID NOS: 10			
		SOFTWARE: Patent in version 3.1			
		SEQ ID NO 2			
		LENGTH: 484			
		TYPE: PRT			
		ORGANISM: Ochrobactrum anthropi			
		US-10-114-810-2			
		Query Match 35.8%; Score 1068; DB 4; Length 484;			
		Best Local Similarity 47.0%; Pred. No. 4.1e-103;			
		Matches 214; Conservative 80; Mismatches 155; Indels 6; Gaps 4;			
QY	18	LLIRGRIVNDQSFYADVHVHEDGLIKQIGENLIVPGIKTIDAHGLMVLPGGVVHTRL	77		
DB	4	VIRGGTVITADTFKADVLIEGKIVAVGDNL---SGDEVIDASGCYIMPGGIDPHTHLQ	60		
QY	78	MPVLGMPADDFCQGTKAALAGTTTMDLHDVPPDTGVSLLAAYEOWRRERADSAACDYS	137		
DB	61	MPFMGTYSDDDFDTGTAALAGTTTMDLHDVPPDTGVSLLAAYEOWRRERADSAACDYS	119		
QY	138	HVDITRWHESIKERLEALVKEKGVNSFLVPMAYKORCQSDSQMYEIPSIIRDLGALAOV	197		
DB	120	HMAITGMNERTFNEMAEVVK-RGINTFKHFMAYKALMVNDDEMFASFQCAELGAMPLV	178		
QY	198	HAENGDIIVEEQKRILLELIGITGPEGHVLSHPREVEAEAVYRAVTIAKQANCLPLYVTKVM	257		
DB	179	HAENGDIIVAOQLQAKLWAGNDGPEAHAYSRPPEVGEATNRAIMTADQAGVPLYVHVHSC	238		
QY	258	KGAADAIQAQKRGVVVFGPEITASIGTDSHYWKNWAKAAAFVTPSPVNPDPPTADHL	317		
DB	239	EQSHEAIRARQKGRVFGPEPLIQHLLTLDSEYHNRDMDYARRVMSPPFR-DKANQDSL	297		
QY	318	TCLLSGDLQVTSAGHCTTTAAQVGNFALIPGTTNGIERMSMWKEKCVASQMDR	377		
DB	298	WAGLAAGSLQCVATDHCFTTRQKRYGIGNFTKIPNGTGLBERMPVLMRSGVTRGLTP	357		
QY	378	NEFVAVTSTNAAKIFNFPYPRKGRVAVGSDADLVINWPKATKIISAKTHNLNVEYNI	437		
DB	358	NEFVAVTSTNAKILINITYPKGAVLPAGADADLVINDPEATRKVSATQHSIDINVFE	417		
QY	438	ECRGAPAVVISQGRVALEDGKMFVTPGAGRFVPRK	472		
DB	418	ELKGLPKMTLSRGRVAFDKGNVTAEPGDGRFIERE	452		
RESULT 12					
		US-09-734-237B-62			
		Sequence 62, Application US/09734237B			
		Patent No. 6818752			
		GENERAL INFORMATION:			
		APPLICANT: Rozzell, J. David			
		APPLICANT: Bui, Peter			
		TITLE OF INVENTION: SYNTHETIC GENES FOR ENHANCED EXPRESSION			
		FILE REFERENCE: B583:40608			
		CURRENT APPLICATION NUMBER: US/09/734,237B			
		CURRENT FILING DATE: 2000-12-08			
		PRIOR APPLICATION NUMBER: 09/494,921			
		PRIOR FILING DATE: 2000-01-31			
		NUMBER OF SEQ ID NOS: 79			
		SOFTWARE: Patent in version 3.1			
		SEQ ID NO 62			
		LENGTH: 495			
		TYPE: PRT			
		ORGANISM: Pseudomonas putida			
		US-09-734-237B-62			
		Query Match 34.9%; Score 1041; DB 4; Length 495;			
		Best Local Similarity 47.1%; Pred. No. 3e-100;			
		Matches 210; Conservative 72; Mismatches 156; Indels 8; Gaps 3;			
QY	17	LLIRGRIVNDQSFYADVHVHEDGLIKQIGENLIVPGIKTIDAHGLMVLPGGVVHTRL	76		
DB	3	LLIRGATVTHESYRADVLQVGLIRALGNLEPPTDCEILDGSGOYLMPGGIDPHTHM	62		
QY	77	QMPVLGMPADDFCQGTKAALAGTTTMDLHDVPPDTGVSLLAAYEOWRRERADSAACDYS	136		
DB	63	QLPFGMTVASEDPFSGTAAGLAGTTTSIIDFVIEPNRQSLLEAFHTWRGWAQKSA-SDYG	121		
QY	137	LHVDITRWHESIKERLEALVKEKGVNSFLVPMAYKORCQSDSQMYEIPSIIRDLGALAOV	196		
DB	122	FHVAITWMSDEVARENGELVAQGVNSFKHFMAYKNAIMADTLVASPERCLQLGAVPT	181		
QY	197	VHAENGDIIVEEQKRILLELIGITGPEGHVLSHPREVEAEAVYRAVTIAKQANCLPLYVTKVM	256		

182 DB VHAENGELVYHLOKLLAQMTGPEAHPLSRPQVEGEAASRAIRIAETIGTPLYVYVHIS 241  
257 QY SKGAADAIAQKRGVGVVVEPITASIGTDSHYWYKNAKAAAFVTSPPVNDPTTADH 316  
242 DB SREALDEITYARAKQPVYGEVLPCHLLDSDVTRDPDMATAAGYVMSPPFRP-----REH 297  
317 QY LTCL---LSSGDLQVTSAGHCTTTTAQKAVGKONFALIPETNGIEMRMWVEKCVASG 373  
298 DB QEALWRGLQSGNLHTTATDHCPCBAEQKAMGRDDFSRIPNGTAGIEDRMVAVLWDAGVNSG 357  
374 QY KMDENEFVAVTSTNAAKIFNFPYPRKGRVAVGSDADLVINWPKATKIISAKTHNLNVEYNI 433  
358 DB RLSMHEFVALTSTNTAKIFNLPKRGAIIRVGADADLVMDPQGTRTLISAQTHHQRVDFNI 417  
434 QY FEGVCEGAPAVVISQGRVALEDGKM 459  
418 DB FEGRTVGVPSHTISQGRVWADGDL 443

RESULT 13  
US-09-734-237B-64  
; Sequence 64 Application US/09734237B  
; Patent No. 6818752  
; GENERAL INFORMATION:  
; APPLICANT: Rozzell, J. David  
; APPLICANT: Bui, Peter  
; APPLICANT: Hua, Ling  
; TITLE OF INVENTION: SYNTHETIC GENES FOR ENHANCED EXPRESSION  
; FILE REFERENCE: B583:40608  
; CURRENT APPLICATION NUMBER: US/09/734,237B  
; CURRENT FILING DATE: 2000-12-08  
; PRIOR FILING DATE: 09/494,921  
; PRIOR FILING DATE: 2000-01-31  
; NUMBER OF SEQ ID NOS: 79  
; SOFTWARE: Patent in version 3.1  
; SEQ ID NO 64  
; LENGTH: 496  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURES:  
; OTHER INFORMATION: Synthetic protein derived from Pseudomonas putida hydantoinase, h  
; OTHER INFORMATION: aving a glycine residue inserted after the initiating methionine  
US-09-734-237B-64

Query Match 34.9%; Score 1041; DB 4; Length 496;  
Best Local Similarity 47.1%; Pred. No. 3e-100;  
Matches 210; Conservative 72; Mismatches 156; Indels 8; Gaps 3;

17 QY LLIRGRVNDQSFYADVHVEDGLIKQIGENLIVPGGIKTIDAHGLMVLPGGVVDVHTRL 76  
4 DB LLIRGATVVTBESSYPADVLCDGLIRAGNLEPPPTDCEILDGSGQYLMPPGIDPHTM 63  
77 QY QMPVLGTPADDDFCQGTKAALAGTTMILDHVPDPTGVSLAAYEOWRERADSAACDYS 136  
64 DB QLPFGMTVASDEDFPGTAAGLAGGTSIIDFVFPNQQSLLEAFHTWRGWAQKSA-SDYG 122  
137 QY LHVDITRWHSIKEELEALYKKGVSFLVPMAYKORCQSDSQMVEIFSIIRDGLAQ 196  
123 DB FVAITWSEQVAEEMGELVAKGVNSFKHFMAYKNAIMAADDTLVASFERCICQLGAVPT 182  
197 QY VHAENGDI VEEOKRLLELIGITGPEGHVLSHPEVEBAEVYAVTIAKQANCLPYTKVM 256  
183 DB VHAENGELVYHLOKLLAQMTGPEAHPLSRPQVEGEAASRAIRIAETIGTPLYVYVHIS 242  
257 QY SKGAADAIAQKRGVGVVVEPITASIGTDSHYWYKNAKAAAFVTSPPVNDPTTADH 316  
243 DB SREALDEITYARAKQPVYGEVLPCHLLDSDVTRDPDMATAAGYVMSPPFRP-----REH 298  
317 QY LTCL---LSSGDLQVTSAGHCTTTTAQKAVGKONFALIPETNGIEMRMWVEKCVASG 373  
299 DB QEALWRGLQSGNLHTTATDHCPCBAEQKAMGRDDFSRIPNGTAGIEDRMVAVLWDAGVNSG 358

374 QY KMDENEFVAVTSTNAAKIFNFPYPRKGRVAVGSDADLVINWPKATKIISAKTHNLNVEYNI 433  
DB RLSMHEFVALTSTNTAKIFNLPKRGAIIRVGADADLVMDPQGTRTLISAQTHHQRVDFNI 418  
434 QY FEGVCEGAPAVVISQGRVALEDGKM 459  
DB FEGRTVGVPSHTISQGRVWADGDL 444

RESULT 14  
US-08-289-709-1  
; Sequence 1 Application US/08289709  
; Patent No. 5523224  
; GENERAL INFORMATION:  
; APPLICANT: Butscher, Helmut; Lang, Gunter; Popp, Friedrich  
; TITLE OF INVENTION: Recombinant D-Hydantoinase, A Process For the  
; TITLE OF INVENTION: Production and Use  
; NUMBER OF SEQUENCES: 4  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Felfe & Lynch  
; STREET: 805 Third Avenue  
; CITY: New York City  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10022  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 5.25 inch, 360 kb storage  
; COMPUTER: IBM PS/2  
; OPERATING SYSTEM: PC-DOS  
; SOFTWARE: Wordperfect  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/289,709  
; FILING DATE: 12-AUGUST-1994  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: P 43 28 829.4  
; FILING DATE: 27-AUGUST-1993  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Hanson, No. 5523224man D.  
; REGISTRATION NUMBER: 30,946  
; REFERENCE/DOCKET NUMBER: BOER 1041  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 688-9200  
; TELEFAX: (212) 838-3884  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 460 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
US-08-289-709-1

Query Match 27.7%; Score 827; DB 1; Length 460;  
Best Local Similarity 39.7%; Pred. No. 8.9e-78;  
Matches 182; Conservative 79; Mismatches 193; Indels 4; Gaps 4;

18 QY LLIRGRVNDQSFYADVHVEDGLIKQIGENLIVPGGIKTIDAHGLMVLPGGVVDVHTRLQ 77  
4 DB IIRNGTIVTATDYEDALLIKDKKIAMICQHIL-EKGAENVIDAKGCYVFPGGIDSHTHLD 62  
78 QY MPVLGTPADDDFCQGTKAALAGTTMILDHVPDPTGVSLAAYEOWRERADSAACDYS 137  
63 DB MPFGGTVTKDDPESGTIAAAGGTTIIDPCLTNKGEPLKKALETWHNKAQKAVIDYGF 122  
138 QY HVDITRWHSIKEELEALYKKGVSFLVPMAYKORCQSDSQMVEIFSIIRDGLAQV 197  
123 DB HLMISEITDDVLEELPKVIAEBGITSFKVPMAYKNVFPQADDGTLVYRTLVAAKELGALVMV 182  
198 QY HAENGDI VEEOKRLLELIGITGPEGHVLSHPEVEBAEVYAVTIAKQANCLPYTKVM 257  
183 DB HAENGDI VYLLTKKALAEKNTBPIYHALTRPPEVEGEATGRACQLTELAGSGLYVHVTC 242  
258 QY KGAADAIAQKRGVGVVVEPITASIGTDSHYWYKNAKAAAFVTSPPVNDPTTADH 317

Db 243 AQAVERIAQARNKGLDVMGECTCQYLVLDSQ-YLEKPDPEGAKYVWSPPLR-EKWHQEV L 300  
QY 318 TCLSSGDLQVTSAGHCTFT-TAQAQVGNKDNFALIPGNTNGIIEERMSMWKCVASGKMD 376  
Db 301 WNALKNGQLQTLGSDQCSDFPKQKELGRGDTKIPNGGPMVEDRVSILFSEGKKGRT 360  
QY 377 ENFVAVTSTNAAKIFNFPYPRKGRVAGSDADLVINPKATKIISAKTHNLNVEYNIFEG 436  
Db 361 LNOFVDMSTRIAKLFGLPFRKGTIAVGSADADLVIFDPIERVISAETHHMAVDYNAPEG 420  
QY 437 VECRGAPAVVISQGRVALEDGKMFVTPGAGRFVPRKTF 474  
Db 421 MKVTGEPVSVLCRGEFVVRDKQFVGKPGYQYLKAGCF 458

## RESULT 15

US-08-602-656-1  
Sequence 1, Application US/08602656  
Patent No. 5679571  
GENERAL INFORMATION:  
APPLICANT: Burtcher, Helmut; Lang, Gunter; Popp, Friedrich  
TITLE OF INVENTION: Recombinant D-Hydantoinase, A Process For the  
NUMBER OF SEQUENCES: 4  
CORRESPONDENCE ADDRESS:  
STREET: 805 Third Avenue  
CITY: New York City  
STATE: New York  
COUNTRY: USA  
ZIP: 10022

COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 5.25 inch, 360 kb storage

COMPUTER: IBM PS/2

OPERATING SYSTEM: PC-DOS

SOFTWARE: Wordperfect

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/602,656

FILING DATE: 16-FEB-1996

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/289,709

FILING DATE: 12-AUGUST-1994

APPLICATION NUMBER: P 43 28 829.4

FILING DATE: 27-AUGUST-1993

ATTORNEY/AGENT INFORMATION:

NAME: Hanson, No. 5679571man D.

REGISTRATION NUMBER: 30,946

REFERENCE/DOCKET NUMBER: BOER 1041

TELECOMMUNICATION INFORMATION:

TELEPHONE: (212) 688-9200

TELEFAX: (212) 838-3884

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 460 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

US-08-602-656-1

Query Match 27.7%; Score 827; DB 1; Length 460;

Best Local Similarity 39.7%; Pred. No. 8.9e-78;

Matches 182; Conservative 79; Mismatches 193; Indels 4; Gaps 4;

QY 18 LIRGRIVNDQSFYADVHVHVEDGLIKQIGENLIVPGGIKTIDAHGLMVLFGGVVDVHRLQ 77

Db 4 IIKNGTIVTATDIYEADLLIKDKGIAMIGQHL-EKGAVIDAKGCVVFPGGIDSHTHLD 62

QY 78 MPVLGWTTPADDFCGTYKAALAGGTTMILDHVPDPTGVSLAAYEQWRERADSAACCDYSL 137

Db 63 MPFGGTVTKDDFBSGTIAAFAFGGTTTIIDFCLTNKGEPKKAJETWHNKAKGRAVIDYGF 122

QY 138 HVDITRWHEISKEELEALVKEKGVNSFLVFMAYKDRCCSDSQMSQYEIFSIIRDLAGALAQV 197  
Db 123 HLMISEITDDVLEELPKVIAEEGITSFKVFMAYKVFQADDGTLRTLVAAKELGALVMV 182  
QY 198 HAENGDIVEEBQKRLLELGITQPEGHVLSHPBEVEABAVYRAVTIAKQANCPLYYTKVMS 257  
Db 183 HAENGDDVLDYLTKKALAEAGNTEPIVHALTRPPEVEGEATGRACQLTELAGSQLYVVHVTC 242  
QY 258 KGAADAIQAQKRGVVVFGPEPITASLGTDSHYWKNWAKAAAFVTSPPVNPDPPTADHL 317  
Db 243 AQAVERIAQARNKGLDVMGECTCQYLVLDSQ-YLEKPDPEGAKYVWSPPLR-EKWHQEV L 300  
QY 318 TCLSSGDLQVTSAGHCTFT-TAQAQVGNKDNFALIPGNTNGIIEERMSMWKCVASGKMD 376  
Db 301 WNALKNGQLQTLGSDQCSDFPKQKELGRGDTKIPNGGPMVEDRVSILFSEGKKGRT 360  
QY 377 ENFVAVTSTNAAKIFNFPYPRKGRVAGSDADLVINPKATKIISAKTHNLNVEYNIFEG 436  
Db 361 LNOFVDMSTRIAKLFGLPFRKGTIAVGSADADLVIFDPIERVISAETHHMAVDYNAPEG 420  
QY 437 VECRGAPAVVISQGRVALEDGKMFVTPGAGRFVPRKTF 474  
Db 421 MKVTGEPVSVLCRGEFVVRDKQFVGKPGYQYLKAGCF 458

Search completed: September 24, 2005, 17:40:45

Job time : 61 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 24, 2005, 17:34:57 ; Search time 113 Seconds  
(without alignments)  
2060.274 Million cell updates/sec

Title: US-09-367-496C-8

Perfect score: 2982  
Sequence: 1 MSFQKKSIPRTSDRLLR.....RTAQKIMAPPGRSNTTSL 572

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1826521 segs, 407012169 residues

Total number of hits satisfying chosen parameters: 1826521

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : Published Applications AA:

- 1: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep.\*
- 2: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB.pep.\*
- 3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB.pep.\*
- 4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pep.\*
- 5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pep.\*
- 6: /cgn2\_6/ptodata/2/pubpaa/PCTUS\_PUBCOMB.pep.\*
- 7: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pep.\*
- 8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep.\*
- 9: /cgn2\_6/ptodata/2/pubpaa/US09A\_PUBCOMB.pep.\*
- 10: /cgn2\_6/ptodata/2/pubpaa/US09B\_PUBCOMB.pep.\*
- 11: /cgn2\_6/ptodata/2/pubpaa/US09C\_PUBCOMB.pep.\*
- 12: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB.pep.\*
- 13: /cgn2\_6/ptodata/2/pubpaa/US10A\_PUBCOMB.pep.\*
- 14: /cgn2\_6/ptodata/2/pubpaa/US10B\_PUBCOMB.pep.\*
- 15: /cgn2\_6/ptodata/2/pubpaa/US10C\_PUBCOMB.pep.\*
- 16: /cgn2\_6/ptodata/2/pubpaa/US10D\_PUBCOMB.pep.\*
- 17: /cgn2\_6/ptodata/2/pubpaa/US10E\_PUBCOMB.pep.\*
- 18: /cgn2\_6/ptodata/2/pubpaa/US10F\_PUBCOMB.pep.\*
- 19: /cgn2\_6/ptodata/2/pubpaa/US11A\_PUBCOMB.pep.\*
- 20: /cgn2\_6/ptodata/2/pubpaa/US11\_NEW\_PUB.pep.\*
- 21: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pep.\*
- 22: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2978	99.9	572	16	US-10-723-860-1487
2	2978	99.9	572	18	US-10-287-436A-1248
3	2978	99.9	572	9	US-09-986-632-10
4	2976	99.8	559	16	US-10-618-281-74
5	2917	97.8	559	17	US-10-498-788-18
6	2628	88.1	512	17	US-10-498-788-18
7	2345	78.6	572	9	US-09-986-632-4
8	2345	78.6	572	15	US-10-205-331-52
9	2345	78.6	572	16	US-10-788-792-128
10	2345	78.6	572	16	US-10-733-969A-95
11	2345	78.6	572	16	US-10-733-969A-100

12	2345	78.6	572	16	US-10-851-921-6	Sequence 6, Appli
13	2192	73.5	570	18	US-10-287-436A-485	Sequence 485, App
14	2192	73.5	570	18	US-10-287-436A-1181	Sequence 181, Ap
15	2184	73.2	570	9	US-09-986-632-6	Sequence 6, Appli
16	2157	72.3	572	9	US-09-986-632-8	Sequence 8, Appli
17	2157	72.3	572	9	US-09-736-457-1815	Sequence 1815, Ap
18	2157	72.3	572	9	US-09-902-941-1815	Sequence 1815, Ap
19	2157	72.3	572	9	US-09-849-626-1815	Sequence 1815, Ap
20	2157	72.3	572	14	US-10-017-754-1815	Sequence 1815, Ap
21	2157	72.3	572	14	US-10-180-198-2	Sequence 2, Appli
22	2157	72.3	572	14	US-10-113-872-1815	Sequence 1815, Ap
23	2157	72.3	572	15	US-10-283-017-1815	Sequence 1815, Ap
24	2157	72.3	572	17	US-10-852-335A-110	Sequence 110, App
25	2101	70.5	573	16	US-10-618-281-70	Sequence 70, Appl
26	1934.5	64.9	532	16	US-10-408-765A-1676	Sequence 1676, Ap
27	1556	52.2	519	16	US-10-408-765A-410	Sequence 410, App
28	1469.5	49.3	564	9	US-09-986-632-2	Sequence 2, Appli
29	1469.5	49.3	564	14	US-10-220-042-2	Sequence 2, Appli
30	1469.5	49.3	564	18	US-10-220-335-204	Sequence 204, App
31	1443	48.4	579	16	US-10-618-281-78	Sequence 78, Appl
32	1208.5	40.5	638	18	US-10-450-763-50979	Sequence 50979, A
33	1071	35.9	539	16	US-10-437-963-118365	Sequence 118365,
34	1068	35.8	484	14	US-10-114-810-2	Sequence 2, Appli
35	1041	34.9	495	10	US-09-734-237B-62	Sequence 62, Appl
36	1041	34.9	495	18	US-10-989-488A-62	Sequence 62, Appl
37	1041	34.9	496	10	US-09-734-237B-64	Sequence 64, Appl
38	1041	34.9	496	18	US-10-989-488A-64	Sequence 64, Appl
39	905	30.3	268	18	US-10-504-582-156	Sequence 156, App
40	840	28.2	332	15	US-10-220-120-214	Sequence 214, App
41	826	27.7	485	15	US-10-424-599-244709	Sequence 244709,
42	819.5	27.5	332	16	US-10-363-829-337	Sequence 337, App
43	809	27.1	461	9	US-09-836-992-1	Sequence 1, Appli
44	725.5	24.3	438	14	US-10-176-584A-2	Sequence 2, Appli
45	714.5	24.0	462	14	US-10-190-471-2	Sequence 2, Appli

ALIGNMENTS

RESULT 1

US-10-723-860-1487  
; Sequence 1487, Application US/10723860  
; Publication No. US20040253606A1  
; GENERAL INFORMATION:  
; APPLICANT: Aziz, Natasha  
; APPLICANT: Gineburg, Wendy M.  
; APPLICANT: Zlotnik, Albert  
; TITLE OF INVENTION: Methods of Diagnosis of Soft Tissue Sarcoma, Compositions &  
; FILE REFERENCE: 05882.0193.NPUS01  
; CURRENT APPLICATION NUMBER: US/10/723,860  
; CURRENT FILING DATE: 2003-11-26  
; PRIOR APPLICATION NUMBER: 60/429,739  
; PRIOR FILING DATE: 2002-11-26  
; NUMBER OF SEQ ID NOS: 8393  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 1487  
; LENGTH: 572  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-723-860-1487

Query Match	99.9%	Score	2978	DB	16	Length	572
Best Local Similarity	99.8%	Pred.	No. 2.4e-261				
Matches	571	Conservative	1	Mismatches	0	Indels	0
Gaps	0						
QY	1	MSFQKKSIPRTSDRLLRGRINVDQSFYADVHVDGLIKQIGENLIVPGIKTIDA	60				
Db	1	MSFQKKSIPRTSDRLLRGRINVDQSFYADVHVDGLIKQIGENLIVPGIKTIDA	60				
QY	61	HGLMVLPGGVVHTRIQMPVLGMPADDFCOQGTAAALAGTTMILDHVPDPTGVSLAAY	120				
Db	61	HGLMVLPGGVVHTRIQMPVLGMPADDFCOQGTAAALAGTTMILDHVPDPTGVSLAAY	120				



Db 301 FVTSPVPPNPPTTADHLTCLSSGDLQVTSAGHCTFTTAQKAVGKONFALIPGTINGIEE 360  
Qy 361 RMSMWKEKCVASGMDNEFVAVTSTNAAKIFNFPYPRKGRVAVGSDADLVINWPKATKII 420  
Db 361 RMSMWKEKCVASGMDNEFVAVTSTNAAKIFNFPYPRKGRVAVGSDADLVINWPKATKII 420  
Qy 421 SAKTHNLNVEYNIPEGVECRGAPAVVISQGRVALEDGKMFVTPGAGRFVPRKTFPPDFVYK 480  
Db 421 SAKTHNLNVEYNIPEGVECRGAPAVVISQGRVALEDGKMFVTPGAGRFVPRKTFPPDFVYK 480  
Qy 481 RIKARNRLAEIHGVPRLYDGPVHEVMVPAKPGSGAPARASCPGKISVPPVRNLHOSGFS 540  
Db 481 RIKARNRLAEIHGVPRLYDGPVHEVMVPAKPGSGAPARASCPGKISVPPVRNLHOSGFS 540  
Qy 541 LSGSQADHIIARRTAQKIMAPPGGRSNTISLS 572  
Db 541 LSGSQADHIIARRTAQKIMAPPGGRSNTISLS 572

## RESULT 4

US-09-986-632-10  
; Sequence 10, Application US/09986632  
; Patent No. US20020119944A1  
; GENERAL INFORMATION:  
; APPLICANT: AGUERA, Michelle  
; TITLE OF INVENTION: Modulation of Ulp1/CRMP activity for the prevention or  
; FILE REFERENCE: P06974US01/BAS  
; CURRENT APPLICATION NUMBER: US/09/986,632  
; CURRENT FILING DATE: 2001-11-09  
; PRIOR APPLICATION NUMBER: US 60/246,751  
; PRIOR FILING DATE: 2000-11-09  
; NUMBER OF SEQ ID NOS: 30  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 10  
; LENGTH: 572  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-986-632-10

Query Match 99.8%; Score 2976; DB 9; Length 572;  
Best Local Similarity 99.8%; Pred. No. 3.7e-261;  
Matches 571; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
Qy 1 MSFGKKSIPIRTSDRLIRGGRIVNDQSFYADVHVEDGLIKQIGENLIVPGGIKTIDA 60  
Db 1 MSFGKKSIPIRTSDRLIRGGRIVNDQSFYADVHVEDGLIKQIGENLIVPGGIKTIDA 60  
Qy 61 HGLMVLPGGVVDVHTRLQMPVLGMPADDFCQGTKAALAGGTTMILDHVFPDGTGVSLLAAY 120  
Db 61 HGLMVLPGGVVDVHTRLQMPVLGMPADDFCQGTKAALAGGTTMILDHVFPDGTGVSLLAAY 120  
Qy 121 EQWRERADSAACDYSLSHVDITRWHSIKEELEALVKEKGVNSFLVPMAYKDRQCSDSQ 180  
Db 121 EQWRERADSAACDYSLSHVDITRWHSIKEELEALVKEKGVNSFLVPMAYKDRQCSDSQ 180  
Qy 181 MYEIFSIIIRDLGALAQVHAENGDIIVREEQKRLLELIGITGPEGHVLSHPREVEAEVYRAV 240  
Db 181 MYEIFSIIIRDLGALAQVHAENGDIIVREEQKRLLELIGITGPEGHVLSHPREVEAEVYRAV 240  
Qy 241 TIAQANCPLYVTWKVSKGAADALAQAKRRGVVVFGEPIITASLGTDGSHYWSKNWAKAAA 300  
Db 241 TIAQANCPLYVTWKVSKGAADALAQAKRRGVVVFGEPIITASLGTDGSHYWSKNWAKAAA 300  
Qy 301 FVTSPVPPNPPTTADHLTCLSSGDLQVTSAGHCTFTTAQKAVGKONFALIPGTINGIEE 360  
Db 301 FVTSPVPPNPPTTADHLTCLSSGDLQVTSAGHCTFTTAQKAVGKONFALIPGTINGIEE 360  
Qy 361 RMSMWKEKCVASGMDNEFVAVTSTNAAKIFNFPYPRKGRVAVGSDADLVINWPKATKII 420  
Db 361 RMSMWKEKCVASGMDNEFVAVTSTNAAKIFNFPYPRKGRVAVGSDADLVINWPKATKII 420

Qy 421 SAKTHNLNVEYNIPEGVECRGAPAVVISQGRVALEDGKMFVTPGAGRFVPRKTFPPDFVYK 480  
Db 421 SAKTHNLNVEYNIPEGVECRGAPAVVISQGRVALEDGKMFVTPGAGRFVPRKTFPPDFVYK 480  
Qy 481 RIKARNRLAEIHGVPRLYDGPVHEVMVPAKPGSGAPARASCPGKISVPPVRNLHOSGFS 540  
Db 481 RIKARNRLAEIHGVPRLYDGPVHEVMVPAKPGSGAPARASCPGKISVPPVRNLHOSGFS 540  
Qy 541 LSGSQADHIIARRTAQKIMAPPGGRSNTISLS 572  
Db 541 LSGSQADHIIARRTAQKIMAPPGGRSNTISLS 572

## RESULT 5

US-10-618-281-74  
; Sequence 74, Application US/10618281  
; Publication No. US20040219609A1  
; GENERAL INFORMATION:  
; APPLICANT: Day, Anthony G.  
; APPLICANT: Estell, David A.  
; APPLICANT: Lyons, Eric H.  
; APPLICANT: Yao, Jian  
; TITLE OF INVENTION: Methods for Modulating Proteases Not  
; FILE REFERENCE: GC773-2  
; CURRENT APPLICATION NUMBER: US/10/618,281  
; CURRENT FILING DATE: 2003-07-11  
; PRIOR APPLICATION NUMBER: US 60/395,325  
; PRIOR FILING DATE: 2002-07-12  
; NUMBER OF SEQ ID NOS: 92  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 74  
; LENGTH: 559  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-618-281-74

Query Match 97.8%; Score 2917; DB 16; Length 559;  
Best Local Similarity 100.0%; Pred. No. 8.2e-256;  
Matches 559; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 14 SDRLLIRGGRIVNDQSFYADVHVEDGLIKQIGENLIVPGGIKTIDAHLMLVLPGGVDVH 73  
Db 1 SDRLLIRGGRIVNDQSFYADVHVEDGLIKQIGENLIVPGGIKTIDAHLMLVLPGGVDVH 60  
Qy 74 TRLOMPVLGMPADDFCQGTKAALAGGTTMILDHVFPDGTGVSLLAAYBQWRERADSAACC 133  
Db 61 TRLOMPVLGMPADDFCQGTKAALAGGTTMILDHVFPDGTGVSLLAAYBQWRERADSAACC 120  
Qy 134 DYSLHVDITRWHSIKEELEALVKEKGVNSFLVPMAYKDRQCSDSQMYEIFSIIRDGLA 193  
Db 121 DYSLHVDITRWHSIKEELEALVKEKGVNSFLVPMAYKDRQCSDSQMYEIFSIIRDGLA 180  
Qy 194 LAQVHAENGDIIVREEQKRLLELIGITGPEGHVLSHPREVEAEVYRAVTTAKQANCPLYVT 253  
Db 181 LAQVHAENGDIIVREEQKRLLELIGITGPEGHVLSHPREVEAEVYRAVTTAKQANCPLYVT 240  
Qy 254 KWSKGAADALAQAKRRGVVVFGEPIITASLGTDGSHYWSKNWAKAAAFAVTSPPVNPDPPTT 313  
Db 241 KWSKGAADALAQAKRRGVVVFGEPIITASLGTDGSHYWSKNWAKAAAFAVTSPPVNPDPPTT 300  
Qy 314 ADHLTCLSSGDLQVTSAGHCTFTTAQKAVGKONFALIPGTINGIERMSVMWKEKCVASG 373  
Db 301 ADHLTCLSSGDLQVTSAGHCTFTTAQKAVGKONFALIPGTINGIERMSVMWKEKCVASG 360  
Qy 374 KMDNEFVAVTSTNAAKIFNFPYPRKGRVAVGSDADLVINWPKATKIIISAKTHNLNVEYNI 433  
Db 361 KMDNEFVAVTSTNAAKIFNFPYPRKGRVAVGSDADLVINWPKATKIIISAKTHNLNVEYNI 420  
Qy 434 FEGVECRGAPAVVISQGRVALEDGKMFVTPGAGRFVPRKTFPPDFVYKRIKARNRLAEIHG 493  
Db 421 FEGVECRGAPAVVISQGRVALEDGKMFVTPGAGRFVPRKTFPPDFVYKRIKARNRLAEIHG 480

QY 494 VPRGLYDGPVHEVWPAKPGSGAPARASCPGKISVPPVRLNHOSGFSLSGQADDDHIARR 553  
 |||||  
 Db 481 VPRGLYDGPVHEVWPAKPGSGAPARASCPGKISVPPVRLNHOSGFSLSGQADDDHIARR 540  
 |||||  
 QY 554 TAQIMAPPGGRSNTSLS 572  
 |||||  
 Db 541 TAQIMAPPGGRSNTSLS 559

## RESULT 6

US-10-498-788-18

; Sequence 18, Application US/10498788

; Publication No. US20050118594A1

; GENERAL INFORMATION:

; APPLICANT: INCYTE GENOMICS, INC.

; APPLICANT: Chawla, Narinder K.

; APPLICANT: Lee, Soo Yeun

; APPLICANT: Ring, Huijun Z.

; APPLICANT: Lee, Ernestine A.

; APPLICANT: Forsythe, Ian J.

; APPLICANT: Khare, Keena

; APPLICANT: Tran, Uyen K.

; APPLICANT: Kable, Amy E.

; APPLICANT: Richardson, Thomas W.

; APPLICANT: Emerling, Brooke M.

; APPLICANT: Lindquist, Erika A.

; APPLICANT: Baughn, Mariah R.

; APPLICANT: Hafalia, April J. A.

; APPLICANT: Jin, Pei

; APPLICANT: Swarnakar, Anita

; APPLICANT: Li, Joana X.

; APPLICANT: Marquis, Joseph P.

; APPLICANT: Lee, Sally

; APPLICANT: Gorvad, Ann E.

; APPLICANT: Sprague, William W.

; APPLICANT: Becha, Shanya D.

; APPLICANT: Elliott, Vicki S.

; TITLE OF INVENTION: ENZYMES

; FILE REFERENCE: PF-1312 PCT

; CURRENT APPLICATION NUMBER: US/10/498,788

; CURRENT FILING DATE: 2004-06-14

; PRIOR APPLICATION NUMBER: US 60/340,357

; PRIOR FILING DATE: 2001-12-14

; PRIOR APPLICATION NUMBER: US 60/342,962

; PRIOR FILING DATE: 2001-12-20

; PRIOR APPLICATION NUMBER: US 60/343,558

; PRIOR FILING DATE: 2001-12-21

; PRIOR APPLICATION NUMBER: US 60/351,107

; PRIOR FILING DATE: 2002-01-22

; NUMBER OF SEQ ID NOS: 84

; SOFTWARE: PERL Program

; SEQ ID NO 18

; LENGTH: 512

; TYPE: PRT

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: misc feature

; OTHER INFORMATION: Incyte ID No: 7506139CD1

US-10-498-788-18

Query Match 88.1%; Score 2628; DB 17; Length 512;  
 Best Local Similarity 89.5%; Pred. No. 1.4e-229;  
 Matches 512; Conservative 0; Mismatches 0; Indels 60; Gaps 1;

QY 1 MSFQKKSIPIRITSDRLIRGGRIVNDQSFYADVHVEDGLIKQIGENLIVPGIKTIDA 60  
 |||||  
 Db 1 MSFQKKSIPIRITSDRLIRGGRIVNDQSFYADVHVEDGLIKQIGENLIVPGIKTIDA 60  
 |||||  
 QY 61 HGLMVLPGGVVDVTRLOMPVLGTPADDPCQGTAAALAGTTMILDHVPDPTGVSLLAAY 120  
 |||||  
 Db 61 HGLMVLPGGVVDVTRLOMPVLGTPADDPCQGTAAALAGTTMILDHVPDPTGVSLLAAY 120  
 |||||  
 QY 121 EQRERADSAACDYSILHVDITRWHSIKEELEALVKEKGVNSFLVPMAYKDRCCQSDSQ 180  
 |||||

Db 121 EQRERADSAACDYSILHVDITRWHSIKEELEALVKEKGVNSFLVPMAYKDRCCQSDSQ 180  
 |||||  
 QY 181 MYEIPSIIRDLAGALAQVHAENGDIYEEBQKRLLLELIGITGPEGHVLSHPBEVEAEAYRAV 240  
 |||||  
 Db 181 MYEIPSIIRDLAGALAQVHAENGDIYEEBQKRLLLELIGITGPEGHVLSHPBEVEAEAYRAV 240  
 |||||  
 QY 241 TIAQANCPLYYTKVMSKGAADAIQAQKRGVVVGERITASLGTGDSHYMSKWKAKAA 300  
 |||||  
 Db 241 TIAQANCPLYYTKVMSKGAADAIQAQKRGVVVGERITASLGTGDSHYMSKWKAKAA 300  
 |||||  
 QY 301 FVTSPPVNDPTTADHLTCLSSGDLQVTSAGHCTFTTAQKAVGKONFALIEGTNGIEE 360  
 |||||  
 Db 301 FVTSPPVNDPTTADHLTCLSSGDLQVTSAGHCTFTTAQKAVGKONFALIEGTNGIEE 360  
 |||||  
 QY 361 RMSVWEKCVASGKMDENEFVAVTSTNAAKIENFVPRKGRVAVGSDADLVINPKATKII 420  
 |||||  
 Db 361 RMSVWEKCVASGKMDENEFVAVTSTNAAKIENFVPRKGRVAVGSDADLVINPKATKII 420  
 |||||  
 QY 421 SAKTHNLNVEYNIPEGVCECGAPAVVISQGRVALLDGKMFVTPGAGRFPVPRKTFPPDVYK 480  
 |||||  
 Db 421 SAKTHNLNVEYNIPEGVCECGAPAVVISQGRVALLDGKMFVTPGAGRFPVPRKTFPPDVYK 480  
 |||||  
 QY 481 RIKARNRLAEIHGVPRLYDGPVHEVWPAKPGSGAPARASCPGKISVPPVRLNHOSGFS 540  
 |||||  
 Db 428 -----LAEIHGVPRLYDGPVHEVWPAKPGSGAPARASCPGKISVPPVRLNHOSGFS 480  
 |||||  
 QY 541 LSGSQADDDHIARRTAQKIMAPPGGRSNTSLS 572  
 |||||  
 Db 481 LSGSQADDDHIARRTAQKIMAPPGGRSNTSLS 512  
 |||||

## RESULT 7

US-09-986-632-4

; Sequence 4, Application US/09986632

; Patent No. US20020119944A1

; GENERAL INFORMATION:

; APPLICANT: AGUERA, Michelle

; TITLE OF INVENTION: Modulation of Uilp/CRMP activity for the prevention or

; FILE REFERENCE: P06974US01/BAS

; CURRENT APPLICATION NUMBER: US/09/986,632

; CURRENT FILING DATE: 2001-11-09

; PRIOR APPLICATION NUMBER: US 60/246,751

; PRIOR FILING DATE: 2000-11-09

; NUMBER OF SEQ ID NOS: 30

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 4

; LENGTH: 572

; TYPE: PRT

; ORGANISM: Homo sapiens

US-09-986-632-4

Query Match 78.6%; Score 2345; DB 9; Length 572;

Best Local Similarity 75.7%; Pred. No. 8.7e-204;

Matches 432; Conservative 72; Mismatches 67; Indels 0; Gaps 0;

QY 1 MSFQKKSIPIRITSDRLIRGGRIVNDQSFYADVHVEDGLIKQIGENLIVPGIKTIDA 60  
 |||||  
 Db 1 MSYQKKNIPIRITSDRLIRGGRIVNDQSFYADYIMEDGLIKQIGENLIVPGSVKTIEA 60  
 |||||  
 QY 61 HGLMVLPGGVVDVTRLOMPVLGTPADDPCQGTAAALAGTTMILDHVPDPTGVSLLAAY 120  
 |||||  
 Db 61 HSRMVIPGGIDVHTPRPDQGMTSADDFQGTAAALAGTTMIDHVVPEPGTSLAFAF 120  
 |||||  
 QY 121 EQRERADSAACDYSILHVDITRWHSIKEELEALVKEKGVNSFLVPMAYKDRCCQSDSQ 180  
 |||||  
 Db 121 DQWRWADSKSCDYSILHVDISEWHKGIQIEEMALVKDHGVNSFLVPMAYKDRFQLTDCQ 180  
 |||||  
 QY 181 MYEIPSIIRDLAGALAQVHAENGDIYEEBQKRLLLELIGITGPEGHVLSHPBEVEAEAYRAV 240  
 |||||  
 Db 181 IVEVLVSIVRDIGALQVHAENGDIYEEBQKRLLLELIGITGPEGHVLSRBEVEAEAVRAI 240  
 |||||



Qy	241	TTAQANCPLYTKWMSKGAADIAQAQRGVVVYPGEBITASLTGDSGHTWSKNWAKAAA	300
Db	241	TTANQNCPLYTKWMSKSSAAVIAQAQRKGTWYVYGEPIITASLTGDSGHTWSKNWAKAAA	300
Qy	301	FVTSPPVNPDPPTTAHLTCLLSSGDLQVTSAGHCTFTTAQAVGKONFALIPETGNGIEE	360
Db	301	FVTSPPLSDPDPTPDLFNLSSCGDLQVTSAGHCTFNTAQAVGKONFTLLIPETNGTGE	360
Qy	361	RMSWMEKCVASGKMDENEFAVATSTNAKIFNYPKRGVAVGSDADLVINWPKATKII	420
Db	361	RMSVIMDKAVTGVKMDENQFVATSTNAKVFNLPRKGRJAVGSDADLVINWDPDSVKTI	420
Qy	421	SAKTHNLVNEYNIPEGVECRGAPAVVVISQGRVALLDGKHFVTPGAGRFVPRKFTFPDVPYK	480
Db	421	SAKTHNSSLEYNIFEGMECRGSPLVVISQGIKVLGDLTHVTEGSGRYIPRKFPDVPYK	480
Qy	481	RIKARNRLAEIHGVRPGLYDGPVHEVMVYPAKPGSGAPARASCCKISVPPVRNLHOSGFS	540
Db	481	RIKARSLRLAELGVFPGLYDGPVCSVSTPKVTVPASSAKTSPAKQAPPVRNLHOSGFS	540
Qy	541	LSGSQADDDHARTTAQKIWAPPGSRNITSLSL 571	
Db	541	LSGAQIIDDNIPIRTTORIVAPPGSRNITSLSL 571	

```

RESULT 8
US-10-205-331-52
; Sequence 52, Application US/10205331
; Publication No. US20040058326A1
; GENERAL INFORMATION:
; APPLICANT: Warner-Lambert Company
; APPLICANT: Lee, Kevin
; APPLICANT: Dixon, Alistair
; APPLICANT: Brooksbank, Robert
; APPLICANT: Pinnock, Robert
; TITLE OF INVENTION: Identification and Use of Molecules Implicated in Pain
; FILE REFERENCE: WL-A-018199
; CURRENT APPLICATION NUMBER: US/10/205,331
; CURRENT FILING DATE: 2002-07-24
; PRIOR APPLICATION NUMBER: GB 0118354.0
; PRIOR FILING DATE: 2001-07-27
; NUMBER OF SEQ ID NOS: 117
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 52
; LENGTH: 572
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Dihydropyrimidinase-related protein
US-10-205-331-52

```

[illegible]

241	Db	TIANQTNCP	LIYITKVMK	SSAARVIAQARK	KGTVVYGE	PITASIGT	DGSHYVSK	NWAKAAA	300
301	Qy	FVTSPPVNP	DDPTTAD	HLITCL	LSGGDLQV	TGSAHCTFTT	QAQKAVG	KDNFALIP	BGTNGIBE 360
301	Db	FVTSPPLS	PDPTTDF	LNLSL	SGDLQV	TGSAHCTFT	QAQKAVG	KDNFTLI	PGTNGTGE 360
361	Qy	RMSVMWEK	CVASGMD	ENBPVAT	STNAKIFN	YPRKGRV	AVGSDADL	VINWPKATKI	420
361	Db	RMSVIMDK	AVVTKMD	ENBPVAT	STNAKVFNL	YPRKGRI	AVGSDADL	VINWPD	SVKTI 420
421	Qy	SAKTHNLN	AEYNI	FE	GVCEGR	GAPVVIS	QGRVALE	DGKMFVTP	PCAGRFVPRKTFPDPVYK 480
421	Db	SAKTHNSS	LEYNIFE	GMBCRG	SPLVVIS	QSKI	VLEB	DGTLHVTE	SGSRYIRPKPFPDPVYK 480
481	Qy	RIKARNLRI	AHICVP	RG	LYDGPV	HEVYVW	PAKPGS	GAPARAS	CPGKISVPPVRNLHQSGFS 540
481	Db	RIKARSLRI	ALRVP	RG	LYDGPV	CEVSVTP	KTVP	PASSAKTIS	PAKQAPVPRNLHQSGFS 540
541	Qy	LSGSQADD	HIARTAK	IMAP	PGGR	SNITSL			571
541	Db	LSGAQID	DDNIP	RETTORI	VAP	PGGR	SNITSL		571

```

RESULT 9
US-10-788-792-128
; Sequence 128, Application US/10788792
; Publication No. US20040191819A1
; GENERAL INFORMATION:
; APPLICANT: Bayer Pharmaceuticals Corporation
; APPLICANT: Eveleigh, Deepa
; APPLICANT: Bigwood, Douglas
; TITLE OF INVENTION: EXPRESSION PROFILES FOR BREAST CANCER AND METHODS OF USE
; FILE REFERENCE: 5152
; CURRENT APPLICATION NUMBER: US/10/788,792
; CURRENT FILING DATE: 2004-02-27
; PRIOR APPLICATION NUMBER: US 60/450,655
; PRIOR FILING DATE: 2003-02-28
; NUMBER OF SEQ ID NOS: 254
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 128
; LENGTH: 572
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-788-792-128

```

Query Match	78.6%	Score 2345;	DB 16;	Length 572;
Best Local Similarity	75.7%;	Pred. No. 8.7e-204;		
Matches	432;	Conservative 72;	Mismatches 67;	Indels 0; Gaps 0;
Qy	1	MSFOGKKSIPRTSDRLIIIRGGIRVNDGDSFYADVHVEDGLIKQIGENLIVPGGIKTIDA	60	
Db	1	MSYQGKKIPRTSDRLIIIRGGIKVNDGDSFYADIYMEDGLIKQIGENLIVPGGVKTIEA	60	
Qy	61	HGLWLVPGGVVDHTRIQMPVLGTPADDFCQGTKAALAGGTTMILDHVPDPDTCVSLAAY	120	
Db	61	HSEWVTPGGIDVHTRIQMPDQGTSADDFFQGTKAALAGGTTMIDHVPPEPCTSLAAF	120	
Qy	121	EQWRPADSAACDYSLSLHVDITRWHSIKEELVKEKGVNSFLVFMAYKORCQCSDSQ	180	
Db	121	DQWRWADSKSCDYSLSLHVDISEWHKGIQEEMEALVKDHGVNSFLVFMAYKORFQLTDCQ	180	
Qy	181	MYEIPGIIIDLGALAOVHAENGDIIVEEQKRLLELGIITGPEGHVLSHPBEVEAEAVYRAV	240	
Db	181	IYEVLSVINDIGALAOVHAENGDIITAEEOQRILDLGITGPEGHVLSRBPBEVEAEAVNRAI	240	
Qy	241	TIYAKQANCPLYITKVMKSGAADAIAQAKRGVVVVFCEPITASLGTGDSHYWSKNWAQAAA	300	
Db	241	TIANTQNCPLYITKVMKSSSAEYIAQARKKGTVVYCEPITASLGTGDSHYWSKNWAQAAA	300	
Qy	301	FVTSPPVNDPPTADHLTCLLSGDLQVTSASHCTTTTAQKAVGKNFALIPGTINGIEE	360	
Db	301	FVTSPPLSPPDPTTPDPLNSLLSCGDIQVTSASHCTTACAKVGKNFTLIPGTINGTEE	360	





## RESULT 14

US-10-287-436A-1181  
; Sequence 1181, Application US/10287436A  
; Publication No. US2005020421A1  
; GENERAL INFORMATION:  
; APPLICANT: CHILDREN'S HOSPITAL MEDICAL CENTER  
; TITLE OF INVENTION: METHOD FOR DIAGNOSIS AND TREATMENT OF  
; TITLE OF INVENTION: RHEUMATOID ARTHRITIS  
; FILE REFERENCE: 10872.514696  
; CURRENT APPLICATION NUMBER: US/10/287,436A  
; PRIOR FILING DATE: 2002-10-31  
; PRIOR APPLICATION NUMBER: US 60/336,220  
; NUMBER OF SEQ ID NOS: 1446  
; SOFTWARE: PASTSEQ for Windows Version 4.0  
; SEQ ID NO 1181  
; LENGTH: 570  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-287-436A-1181

Query Match 73.5%; Score 2192; DB 18; Length 570;  
Best Local Similarity 70.1%; Pred. No. 7.1e-190;  
Matches 401; Conservative 91; Mismatches 78; Indels 2; Gaps 2;

QY	1	MSFOGKKSIPRITSDRLIRGGRIVNDQSFYADVHVHVEDGLIKQIGENLIVPGGIKTIDA	60
DB	1	MSYQGGKNIPRITSDRLIRGGRIVNDQSFYADVHVHVEDGLIKQIGENLIVPGGVKTIEA	60
QY	61	HGLMWLPGGVDVHTRLQMPVLGTPADDFCQGTKAALAGGTTMILDHVPDPDTGVSLLAAY	120
DB	61	NGKMWIPGGIDVHTHFQMPYKGMTVDDFFQGTKAALAGGTTMILDHVPPESSILEAY	120
QY	121	EQWRERADSAACDYSILHVDITRWHSIEKEELBALVKEGVNSFLVFMAYKORCQSDSQ	180
DB	121	EKWREWADGKSCCDYALHVDITHWNDSVKQEVQNLIKDKGVNSFMVYMAKOLYQVSNTE	180
QY	181	MYEIPSIIRDLGALAOVHAENGDIIVEEOKRLELIGITGPEGHVLSHPPEVEAEAYRAV	240
DB	181	LYEIFTCLEGALAOVHAENGDIITAOEQTRMLEMIGITGPEGHVLSRPELEAEAVRAI	240
QY	241	TIQAQNCPLVYTKVMSKGAADAIQAQRGVVVFGEPTASLGTGSHYWSKNWAKAAA	300
DB	241	TIASQTNCPVYTKVMSKSAADLISQARKKGNVVFGEPTASLGTGSHYWSKNWAKAAA	300
QY	301	FVTSPPVNPDPPTADHLTCLSSGDLQVTSAGHCTFTTAQKAVGKONFALIPGTINGIEE	360
DB	301	FVTSPLSPDPTPDYINSLASGDLQLSGSAHCTFSTAQAIGKONFTAIPGTINGVEE	360
QY	361	RMSMWKCVASGKMDENEFVAVTSTNAAKIFNFPYPRKGRVAVGSDADLVINPKATKII	420
DB	361	RMSVIWDKAVATGKMDENGFVAVTSTNAAKIFNLYPRKGRISVGSDDLVIWDPDAVKIV	420
QY	421	SAKTHNLNVEYNI FEGVECRGAPAVVISQGRVALEDGKMFVTPGAGRFVPRKTFPDPVYK	480
DB	421	SAKNHQAAYNI FEGMELRGAPLVVICQKIMLEDGNLHVTCGAGRFPCSPFSDVYK	480
QY	481	RIKARNRLAIEHGVPRGLYDGPVHEVMVPAKPGSGAPARASCPGKISVPPVRLNHQSGFS	540
DB	481	RIKARKKMDLHVAVRGMYDGPVFDLTTPPKGTPAGSARGSPTRPN-PPVRLNHQSGFS	539
QY	541	LSGSQADDDHARTAKINAPGGRSNTLS	572
DB	540	LSGTQVDEGV-RSASKRIIVAPGGRSNTLS	570

## RESULT 15

US-09-986-632-6  
; Sequence 6, Application US/09986632  
; Patent No. US20020119944A1  
; GENERAL INFORMATION:  
; APPLICANT: AGUERA, Michelle

; TITLE OF INVENTION: Modulation of Ulip/CRMP activity for the prevention or  
; TITLE OF INVENTION: treatment of myelin disorders  
; FILE REFERENCE: P06974US01/BAS  
; CURRENT APPLICATION NUMBER: US/09/986,632  
; CURRENT FILING DATE: 2001-11-09  
; PRIOR APPLICATION NUMBER: US 60/246,751  
; PRIOR FILING DATE: 2000-11-09  
; NUMBER OF SEQ ID NOS: 30  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 6  
; LENGTH: 570  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-986-632-6

Query Match 73.2%; Score 2184; DB 9; Length 570;  
Best Local Similarity 69.8%; Pred. No. 3.8e-189;  
Matches 399; Conservative 92; Mismatches 79; Indels 2; Gaps 2;

QY	1	MSFOGKKSIPRITSDRLIRGGRIVNDQSFYADVHVHVEDGLIKQIGENLIVPGGIKTIDA	60
DB	1	MSYQGGKNIPRITSDRLIRGGRIVNDQSFYADVHVHVEDGLIKQIGENLIVPGGVKTIEA	60
QY	61	HGLMWLPGGVDVHTRLQMPVLGTPADDFCQGTKAALAGGTTMILDHVPDPDTGVSLLAAY	120
DB	61	NGKMWIPGGIDVHTHFQMPYKGMTVDDFFQGTKAALAGGTTMILDHVPPESSILEAY	120
QY	121	EQWRERADSAACDYSILHVDITRWHSIEKEELBALVKEGVNSFLVFMAYKORCQSDSQ	180
DB	121	EKWREWADGKSCCDYALHVDITHWNDSVKQEVQNLIKDKGVNSFMVYMAKOLYQVSNTE	180
QY	181	MYEIPSIIRDLGALAOVHAENGDIIVEEOKRLELIGITGPEGHVLSHPPEVEAEAYRAV	240
DB	181	LYEIFTCLEGALAOVHAENGDIITAOEQTRMLEMIGITGPEGHVLSRPELEAEAVRAI	240
QY	241	TIQAQNCPLVYTKVMSKGAADAIQAQRGVVVFGEPTASLGTGSHYWSKNWAKAAA	300
DB	241	TIASQTNCPVYTKVMSKSAADLISQARKKGNVVFGEPTASLGTGSHYWSKNWAKAAA	300
QY	301	FVTSPPVNPDPPTADHLTCLSSGDLQVTSAGHCTFTTAQKAVGKONFALIPGTINGIEE	360
DB	301	FVTSPLSPDPTPDYINSLASGDLQLSGSAHCTFSTAQAIGKONFTAIPGTINGVEE	360
QY	361	RMSMWKCVASGKMDENEFVAVTSTNAAKIFNFPYPRKGRVAVGSDADLVINPKATKII	420
DB	361	RMSVIWDKAVATGKMDENGFVAVTSTNAAKIFNLYPRKGRISVGSDDLVIWDPDAVKIV	420
QY	421	SAKTHNLNVEYNI FEGVECRGAPAVVISQGRVALEDGKMFVTPGAGRFVPRKTFPDPVYK	480
DB	421	SAKNHQAAYNI FEGMELRGAPLVVICQKIMLEDGNLHVTCGAGRFPCSPFSDVYK	480
QY	481	RIKARNRLAIEHGVPRGLYDGPVHEVMVPAKPGSGAPARASCPGKISVPPVRLNHQSGFS	540
DB	481	RIKARKKMDLHVAVRGMYDGPVFDLTTPPKGTPAGSARGSPTRPN-PPVRLNHQSGFS	539
QY	541	LSGSQADDDHARTAKINAPGGRSNTLS	572
DB	540	LSGTQVDEGV-RSASKRIIVAPGGRSNTLS	570

Search completed: September 24, 2005, 17:50:06  
Job time : 115 secs